

Predictability of Idealized Thunderstorms in Buoyancy–Shear Space

JOHN R. LAWSON

Cooperative Institute for Mesoscale Meteorological Studies, University of Oklahoma, and NOAA/OAR/National Severe Storms Laboratory, Norman, Oklahoma

(Manuscript received 27 July 2018, in final form 24 April 2019)

ABSTRACT

Thunderstorms are difficult to predict because of their small length scale and fast predictability destruction. A cell's predictability is constrained by properties of the flow in which it is embedded (e.g., vertical wind shear), and associated instabilities (e.g., convective available potential energy). To assess how predictability of thunderstorms changes with environment, two groups of 780 idealized simulations (each using a different microphysics scheme) were performed over a range of buoyancy and shear profiles. Results were not sensitive to the scheme chosen. The gradient in diagnostics (updraft speed, storm speed, etc.) across shear–buoyancy phase space represents sensitivity to small changes in initial conditions: a proxy for inherent predictability. Storm evolution is split into two groups, separated by a U-shaped bifurcation in phase space, comprising 1) cells that continue strengthening after 1 h versus 2) those that weaken. Ensemble forecasts in regimes near this bifurcation are hence expected to have larger uncertainty, and adequate dispersion and reliability is essential. Predictability loss takes two forms: (i) chaotic error growth from the largest and most powerful storms, and (ii) tipping points at the U-shaped perimeter of the stronger storms. The former is associated with traditional forecast error between corresponding grid points, and is here counterintuitive; the latter is associated with object-based error, and matches the mental filtering performed by human forecasters for the convective scale.

1. Introduction

The skill of a numerical weather prediction (NWP) forecast is dependent on the flow regime (e.g., Palmer 1988), all else being equal. In chaotic flow, even the smallest change in initial conditions (ICs) can grow quasi exponentially, and ultimately yields a horizon of predictability relative to the scale of motion (Lorenz 1969; Palmer et al. 2014). Locally, predictability is destroyed quickly by moist convection (Zhang et al. 2002), potentially through the mechanism of tipping points (e.g., parcels reaching the level of free convection) that accentuate sensitivity to ICs (the so-called butterfly effect; Lorenz 1995). Furthermore, as limited-area-model forecasts are run at a higher resolution, the ratio of the vertical depth to the horizontal grid spacing of the model becomes larger: the elliptical dimension increases (Schertzer and Lovejoy 1985). In this circumstance, error growth is not confined to a quasi-2D field, and has a relative greater degree of freedom in the vertical. As such, the greater the elliptical dimension,

the more detrimental small errors in time and space become to the forecast (Tennekes 1978).

Nonlinear dynamical systems across many disciplines may demonstrate extreme sensitivity to ICs, giving rise to tipping points (e.g., Gleick 1987; Coffey and Parker 2018). For example, consider an air parcel almost at its level of free convection. At this tipping point, an infinitesimal change in ICs of the parcel's potential temperature or specific humidity is enough to nudge the atmosphere from one regime (no thunderstorm) to another (thunderstorm). Thunderstorm development then triggers phenomena such as gravity waves (Curry and Murty 1974), surface pressure changes (Markowski and Richardson 2010, 218–220), and upscale growth through cell mergers (e.g., Burke and Schultz 2004). The “no thunderstorm–thunderstorm” bifurcation results in a binary choice of solutions, or basins of attraction separated by a sharp divide. This is analogous to standing on a

Publisher's Note: This article was revised on 3 September 2019 to correct a typographical error that switched the two panel labels in Fig. 7. The corresponding labels as identified in the accompanying text were also corrected.

Corresponding author: John R. Lawson, john.lawson@noaa.gov

DOI: 10.1175/JAS-D-18-0218.1

© 2019 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

mountain peak along the Continental Divide of the North American Rockies and pouring water on one side. A small change in position will mean the water will end up in a different ocean. We can likewise conceptualize thunderstorm development in a “landscape” where basins of attraction (like watersheds; Grebogi et al. 1987; Nusse and Yorke 1996) are the thunderstorm’s mode, intensity, or any characteristic; this approach is demonstrated herein.

The application of theoretical predictability to operational NWP has yielded great insight into potential forecast-skill limits: for example, those associated with low-frequency variability (Palmer 1988), extratropical cyclones (Zhang et al. 2002; McMurdie and Ancell 2014), mesoscale convective systems (Wandishin et al. 2008, 2010; Rodwell et al. 2013; Durran and Weyn 2016; Lillo and Parsons 2017), supercells (Cintineo and Stensrud 2013; Flora et al. 2018), tornadogenesis (Zhang et al. 2016; Coffey et al. 2017), cloud-resolving models (Hohenegger et al. 2006), and so on. An estimate of maximum skill for a given phenomenon or regime, while impossible to diagnose precisely, can normalize probabilistic estimates for forecasters. A large ensemble spread in a less predictable regime may not be poor performance from the modeling system itself, but a fair reflection of the uncertainty given observational error. In fact, if forecasters and researchers can identify the most unpredictable scenarios, where the most value is gained from small improvements in IC estimates (e.g., in situ observations, satellites, radar assimilation), targeted observations (e.g., Ancell and Hakim 2007), or stochastic perturbations (e.g., Buizza et al. 1999), forecasts could be substantially improved if the model is nudged into the correct “watershed” through adjustment of the forecast’s trajectory through phase space (as in Leroy and Rodwell 2014; Berner et al. 2017). Of course, the correct watershed is not known a priori, but a sufficient sampling of phase space by the ensemble members is defined by this exploration of potential basins of attraction.

Hence, the variation of this theoretical predictability is sensitive to the large-scale environment. Small scales may or may not inherit extended predictability from larger scales (Durran and Weyn 2016), but large-scale flow does constrain the convective evolution of a thunderstorm (Weisman and Klemp 1982, hereafter WK82). In WK82, the authors tested the sensitivity of thunderstorm evolution to one definition of buoyancy (related to surface mixing ratio q_{s0} through a prescribed vertical profile)¹ and vertical wind shear U (hereby “shear”) in

idealized numerical simulations. Their results revealed a “Goldilocks zone” for maximizing rotation and vertical motion; the nature and strength of the simulated cells changed through different regions of buoyancy–shear phase space. The WK82 methodology was later extended to investigate sensitivity to hodograph shape (Weisman and Klemp 1984; McCaul and Weisman 2001) and the depth of the moist layer (McCaul and Cohen 2002), while Kirkpatrick et al. (2011) measured multiple novel storm characteristics as a function of buoyancy–shear phase space. Furthermore, Potvin and Flora (2015) ran idealized supercell simulations using observed atmospheric profiles from three high-profile tornadic supercell cases, and measured sensitivity to horizontal grid spacing.

The increase in computer power since WK82 now allows greater granularity with which to test sensitivity of thunderstorm evolution to buoyancy and shear. Herein, the methodology of WK82 is extended to construct a “landscape” of attraction basins as a function of atmospheric-profile ICs (outlined in section 2). In section 3, various storm diagnostics are analyzed for their variation in buoyancy–shear (phase) space, and the implications on predictability are synthesized and summarized in sections 4 and 5, respectively.

2. Methodology

a. Numerical model configuration

All simulations were performed with the Weather Research and Forecasting (WRF) Model (Powers et al. 2017), version 3.7, in idealized mode (em_quarter_ss). However, all vertical profiles herein imply a straight hodograph (i.e., Fig. 1), rather than one that is curved (as in the eponymous, default quarter-circle hodograph). Two matrices of runs were generated: one using the NSSL two-moment microphysics (Mansell et al. 2010), and the other with Morrison two moment (Morrison et al. 2009), to test for sensitivity of the results to microphysical scheme. In the absence of a cumulus parameterization, error from the microphysics parameterization is expected to dominate other parameterized processes (Jankov et al. 2005). Two two-moment schemes were chosen because of their increased realism over single-moment schemes (Igel et al. 2015). All other model options were left as default. Herein, we use an otherwise-identical model setup for each microphysics scheme to primarily detect sensitivity to ICs, rather than model error, as is the scope of the present study. However, this is not to diminish the importance of model uncertainty on thunderstorm forecasts (Stensrud et al. 2000; Lawson and Gallus 2016a), or errors stemming from truncation (Berner et al. 2012), numerical diffusion (Knievel et al. 2007), etc.

Idealized WRF runs were fed with a sounding representative of warm-season vertical profiles, in proximity to

¹This definition of buoyancy, as discussed in Markowski and Richardson (2010, 20–21) and used here and in WK82, uses a representative surface parcel while ignoring the impact of hydrometeors, pressure perturbations, and water vapor.

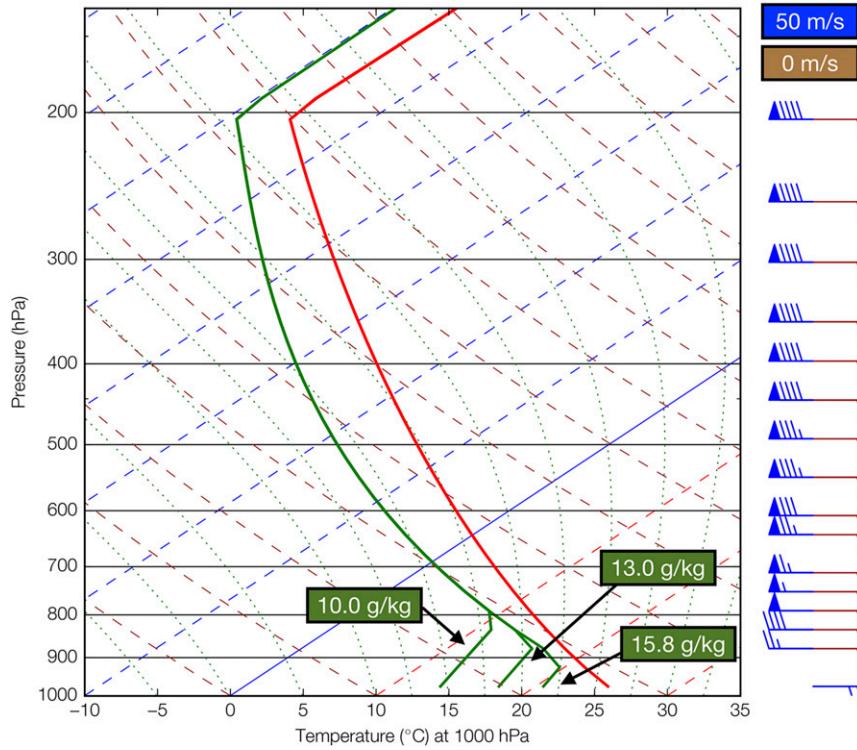


FIG. 1. Skew T - $\log p$ graph showing atmospheric profiles of humidity, temperature, and wind speed and how these variables change across phase space.

supercellular development in the continental United States, and associated hazardous weather (WK82). Figure 1 displays how initial surface water vapor mixing ratio q_{v0} was varied between 10.0 and 16.0 g kg^{-1} , while vertical wind shear U was varied according to Eq. (4) in WK82:

$$U = U_s \tanh\left(\frac{z}{3}\right), \quad (1)$$

where U_s is a factor varied between 0 and 50 m s^{-1} , and is approximately equal to the initial 0–6-km vertical wind shear. The resulting wind profile traces a straight-line hodograph (not shown). Environmental potential temperature $\bar{\theta}$ is varied with height z as in WK82:

$$\bar{\theta}(z) = \begin{cases} \theta_0 + (\theta_{\text{tr}} - \theta_0) \left(\frac{z}{z_{\text{tr}}}\right)^{5/4}, & z \leq z_{\text{tr}} \\ \theta_{\text{tr}} \exp\left(\frac{g}{c_p T_{\text{tr}}}(z - z_{\text{tr}})\right), & z > z_{\text{tr}}, \end{cases} \quad (2)$$

where $c_p = 1004 \text{ J K}^{-1} \text{ kg}^{-1}$ is the specific heat of dry air at constant pressure, $g = 9.81 \text{ m s}^{-2}$ is acceleration under gravity, θ_0 is the surface potential temperature (set to 300 K), and the subscript “tr” denotes constants at the tropopause: height $z_{\text{tr}} = 12 \text{ km}$, potential temperature $\theta_{\text{tr}} = 343 \text{ K}$, and dry-bulb temperature $T_{\text{tr}} = 213 \text{ K}$.

Environmental relative humidity $\overline{\text{RH}}$ is then varied as in WK82:

$$\overline{\text{RH}}(z) = \begin{cases} 1 - \frac{3}{4} \left(\frac{z}{z_{\text{tr}}}\right)^{5/4}, & z \leq z_{\text{tr}} \\ 0.25, & z > z_{\text{tr}}, \end{cases} \quad (3)$$

with the exception of an imposed well-mixed boundary layer, constrained by the constant q_{v0} for each simulation. Given the fixed vertical temperature profile, there is a simple relationship between q_{v0} and convective available potential energy (CAPE), and hence we use the two interchangeably to indicate latent instability (Schultz and Schumacher 1999). Herein, CAPE will refer to lowest-100-hPa mixed-layer CAPE (calculation from Blumberg et al. 2017), and the conversion between CAPE and q_{v0} is shown on the coupled y axes for relevant heat maps below. The author will refer to *buoyancy*, q_{v0} , and CAPE interchangeably, where appropriate, within the framework of the present study. Likewise, the following text refers to U , U_s , and *vertical wind shear* (or simply *shear*) interchangeably, where appropriate. This is not to say these interchangeable terms are equivalent, but rather, they are strongly correlated quantities and the results likely apply similarly to the various measures of

TABLE 1. Definitions of low, middle, high, and extreme levels of CAPE, q_{v0} , and U_s used in the manuscript. The q_{v0} ranges are estimated from the CAPE– q_{v0} relationship (see y axes of Fig. 4, etc.). Informed by U.S. climatology for severe and tornadic weather events in Grams et al. (2012, their Figs. 14a, 15a, 16a, 17a).

Variable	Low	Middle	High	Extreme
U_s (m s^{-1})	0–15	15–25	25–35	35+
CAPE (J kg^{-1})	0–500	500–1500	1500–2500	2500+
q_{v0} (g kg^{-1})	0–10.5	10.4–12.8	12.8–14.8	14.8+

buoyancy and vertical wind shear. The author’s definitions of low, mid-, high, and extreme regimes of CAPE, q_{v0} , and U_s are shown in Table 1, and were informed by a climatology of severe weather environments in the United States (Grams et al. 2012). Herein, we define *severe* near-surface wind as exceeding 25.7 m s^{-1} (50 kt or 58 mph), in concurrence with the Storm Prediction Center (SPC) definitions found online (<https://www.spc.noaa.gov/misc/about.html>; accessed 1 January 2019). (Hail and tornadoes are not explicitly addressed herein.)

Each idealized simulation was run for 2 h, though 50 runs (6%; mainly in the extreme- q_{v0} regime) developed numerical instability in vertical motion while using the Morrison microphysics scheme and crashed early (cf. 51 runs for the NSSL two-moment scheme). These runs are represented by a black square in the following heat-map figures, denoting either missing data or nonsensical values. While these simulations could have been rerun with a smaller time step, this would have required doing so for the entire dataset for consistency. The small fraction of runs that end prematurely is unlikely to affect the general conclusions herein.

The present phase-space exploration differs from WK82 primarily by increasing the granularity within buoyancy–shear space. WK82 used data from 60 simulations with 1.0 g kg^{-1} steps in buoyancy and 5 m s^{-1} steps in shear. Their simulations used 2-km horizontal grid spacing using the Klemp–Wilhelmson model with Chen modifications, and one microphysics parameterization (a Kessler-like warm-rain scheme). In contrast, 780 runs per microphysics scheme were performed herein, with two schemes; environmental profiles were varied in steps of 0.2 g kg^{-1} in q_{v0} (effectively ~ 100 – 150 J kg^{-1} in CAPE) and 2 m s^{-1} in U_s . Horizontal grid spacing of 3 km was chosen to mimic current operational convection-allowing models (e.g., Wheatley et al. 2015), and to strike the balance between a large dataset (considering memory and computational constraints) and sufficiently fine spacing to resolve the fundamental nature of the supercells (Potvin and Flora 2015). Fifty vertical levels between 1000 and 100 hPa were used, $\sim 400 \text{ m}$ apart, to mimic typical operational standards. Otherwise, the configuration was kept constant between

runs, including no Coriolis effect, Rayleigh damping 5 km from the top of the model (option 2 in the WRF namelist configuration), and a free-slip surface (i.e., surface physics were turned off). A mean easterly flow of 5 m s^{-1} was imparted on the initial sounding, such that when combined with a domain size of 120 by 151 grid points (i.e., 360 km by 453 km), no cells exited the open boundaries during the simulations. The long axis of this rectangular domain is oriented along the expected axis of storm movement (i.e., from left to right as viewed from above). The initiation of convection within the idealized WRF runs was via a warm bubble of maximum perturbation temperature equal to 3 K, with a 10-km horizontal radius and 1.5-km vertical radius. This perturbation was then applied to the potential-temperature IC field one-quarter along the x axis and one-half along the y axis (i.e., center left if viewed from above).

b. Predictability and uncertainty

At this point, we differentiate between two concepts of predictability; however, the sense of predictability is not always explicitly stated in the following sections when the implication applies to both senses. First, a point-by-point numerical verification assesses traditional or *gridpoint predictability*, the form primarily discussed by Lorenz (1963). In this framework, predictability is lost at a given point in time when the error between a pair of similar forecasts approaches the climatological error: the so-called predictability horizon (Palmer et al. 2014). Forecast evaluation metrics typically used on the synoptic scale, such as root-mean-square error, verify smooth fields with no explicit tolerance for temporal and spatial phase errors. However, such traditional scores may grossly underestimate the predictability horizon for events on smaller scales (e.g., Reinecke and Durran 2009), as error saturation is reached quickly, and phase errors of objects (e.g., thunderstorms) grow larger as a fraction of the object’s size. The implication of gridpoint error saturation is that all useful prognostic information has been extracted as the horizon is reached, when in fact useful information is still available to the forecaster (contingent on relevant filtering and/or statistical processing, either explicitly, or implicitly performed as part of the forecast process).

In response, recent verification scores and schemes have been developed that include tolerance for temporal and/or spatial phase errors (such as the fractions skill score; Roberts and Lean 2008), or that reduce a heterogeneous field to objects, such as the structure–amplitude–location method (Wernli et al. 2008) or the methodology recently applied in Skinner et al. (2018). These schemes avoid a double penalty in time and space for fields associated with discrete phenomena (Gilleland et al. 2009, 2010), and address a predictability that has been extended

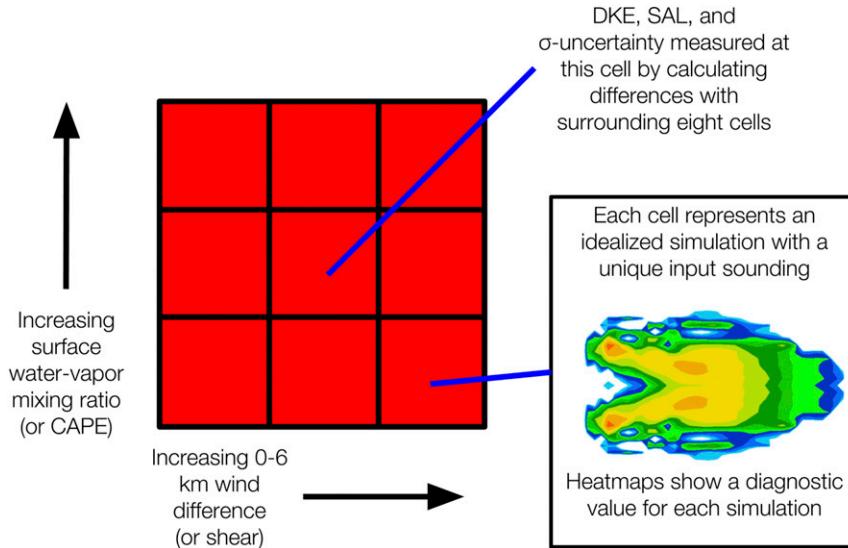


FIG. 2. Schematic diagram outlining the heat map and uncertainty figures in the present paper for a given microphysics parameterization scheme. This methodology is repeated for each nine-pixel window in the phase-space matrix. Before plotting, the σ -uncertainty field is smoothed, or SAL-component standard deviation is computed, as discussed in text.

through filtering (as reviewed in Williams 1997). This is referred to herein as *object-based predictability*.

High predictability (i.e., a longer predictability horizon) is associated with low uncertainty, and vice versa. Uncertainty maxima in buoyancy–shear space highlight regions with increased sensitivity to small changes in ICs (lower predictability); however, given such pessimistic estimates of predictability occur in the absence of error tolerance, the conceptual results are dependent on whether gridpoint or object-based predictability is being measured. To measure traditional predictability, the variation between a number of adjacent cells (Fig. 2) at a given time step is measured by difference kinetic energy (DKE; e.g., Zhang et al. 2002), computed as follows:

$$DKE = 0.5 \sum U'_{ijk}{}^2 + V'_{ijk}{}^2, \quad (4)$$

where U' and V' are the differences in x - and y -direction winds, respectively, computed at all grid points (i, j, k) . This computes the difference at every grid point between every permutation pair within a given group of adjacent simulations; herein, we compute DKE over 3×3 windows. In the following, DKE represents an integrated loss of predictability: the larger the value, the closer to error saturation. A unit-time derivative, hereby termed dDKE, can be considered proportional to the *rate of predictability destruction*, and is shown in the following to assess regions of phase space that are deemed less predictable in the gridpoint sense. DKE was not computed for central cells that lay on the perimeter of phase space (i.e., at a minimum

or maximum value of U_s and q_{v0}), as the 3×3 window partly lies outside of phase space. These edge pixels are represented by black pixels in the heat map.

To complement this energy-based approach, object-based predictability is measured by computing the standard deviation of structure–amplitude–location (SAL) components (Wernli et al. 2008), which has been adapted for radar reflectivity (Lawson and Gallus 2016b). In short, SAL identifies objects in two fields; the two fields are often simulation-versus-observation comparisons, but herein we analyze adjacent idealized simulations only. The SAL method decomposes differences between these fields into three components, first by masking below a given threshold (here, 5 dBZ), and only identifying objects that surpass a pixel threshold (here, five pixels, which is approximately the minimum effective resolution). Results for real-world verification are sensitive to choice of threshold and footprint (Lawson and Gallus 2016b). However, in the present paper, the knowledge that a maximum of one storm can be generated reduces much of the spurious signal (Wernli et al. 2008), and preliminary results were not substantially affected by the choice of settings. The SAL components are

- structure S (between -2 and $+2$), which considers the object size and peakedness/flatness, such that a positive S value for a given simulation may indicate a flatter (i.e., more stratiform) character of composite reflectivity compared to the control;
- amplitude A (between -2 and $+2$), which measures a normalized domainwide accumulation, such that a positive A value indicates a positive reflectivity bias; and

• location L (between 0 and +2), a sum of Euclidean differences between 1) object centers of mass, weighted by their integrated quantity, and 2) domainwide center of mass. Hence, a larger L component represents a larger distance error between identified objects. No object matching is required, in contrast to the methodology of Skinner et al. (2018) and others.

SAL is deployed herein to evaluate reflectivity at 1 km above ground level (AGL) between a given simulation and the eight most similar initial-profile simulations (i.e., the eight cells surrounding the given location in the matrix of phase space). As with DKE, the runs around the perimeter of phase space are neglected, as a full nine-pixel computation cannot be performed, and are represented by black pixels in the heat map. After SAL is computed (Fig. 2), the standard deviation of each SAL component's phase space is applied (over the same 3×3 windows, with no smoothing). In contrast to other figures of uncertainty, we show this standard deviation as the diagnostic variable (i.e., the pixel colors), to highlight potential tipping points in phase space. The standard deviation of SAL is used (rather than mean values) to highlight the variation of component differences between a given cell (a control) and its neighbors in phase space. Use of the mean would smooth out cancelling negative and positive errors, and mistakenly emphasize regions of consistently high (but similar types of) error.

The uncertainty of other diagnostic fields, shown in heat-map figures herein, is computed with a moving 3×3 -window standard deviation operation. This standard deviation field is then passed through a two-sigma Gaussian smoother to remove small variations (i.e., noise) on the order of differences between experiments with different microphysics schemes. The resulting field is hereby termed σ uncertainty, where a subscript will indicate the field in question (e.g., σ_{rain} for rainfall diagnostics), and is contoured with black lines in the following heat-map figures. The larger the σ uncertainty, the larger the sensitivity of the thunderstorm to small changes in ICs, and hence the shorter the theoretical predictability horizon. Later sections will discuss whether the regions of large σ uncertainty correspond to gridpoint or object-based predictability.

c. Other diagnostics

As the simulations herein are run at 3-km horizontal grid spacing, little can be said explicitly regarding tornadogenesis and its predictability. However, updraft helicity (UH; Kain et al. 2008) represents a proxy for mesocyclonic rotation near the surface (0–2 km), which in turn provides a potential detection of tornadogenesis (Trapp et al. 2005). A weaker relationship exists between

tornadogenesis and mesocyclone detection at midlevels (2–5 km), and UH computed within this layer is also included here for comparison with low-level UH. Because of the volume of computations, UH is computed by 1) identifying the nearest model levels to the relevant heights AGL and 2) integrating between those levels (inclusive) as such:

$$\text{UH} = \sum_{l=x}^{l=X} \Delta z \bar{\xi} \bar{w}, \quad (5)$$

where l is the model level, x and X are the lower- and upper-level values respectively, $\bar{\xi}$ and \bar{w} are the vertical relative vorticity and vertical motion averaged for the layer, respectively, and Δz is the difference in meters between each model level. Herein, we use the lowest model level $l=0$ (approximately 200 m AGL), $l=6$ (approximately 2 km), and $l=13$ (approximately 5 km). Hence, UH_{02} is approximately the 0–2-km UH using $x=0$ and $X=6$, while UH_{25} is the 2–5-km UH using $x=6$ and $X=13$.

Storm speed is estimated for a given time by first identifying the 500-hPa updraft maximum 10 min before and after the given time. A storm is deemed present if the location of this updraft corresponds to simulated composite reflectivity greater than 25 dBZ. If no storm is detected, this cell is ignored (and represented by a black pixel in heat maps); otherwise, the distance traveled is computed via the Pythagorean relationship, and converted to meters per second.

To estimate storm size within the simulation for a given time, the number of pixels N_c containing RH greater than 99.9% at $l=6$ (~2 km AGL, or 700 hPa) was converted to equivalent storm diameter d_s :

$$d_s = 2 \times 3 \times 2 \times \sqrt{\frac{N_c}{\pi}}, \quad (6)$$

where the first three integer factors account for the half domain, horizontal grid spacing (in km), and conversion from radius to diameter, respectively. Because of the eccentricity of storm objects, the estimated equivalent radius will be smaller than if measured by, for example, a convex hull operation, and this underestimation is likely worse for higher shear values (because of deformation). However, this is unlikely to affect the general conclusions: preliminary work showed similar patterns when total pixel fraction of the half domain was used, or when the RH threshold was lowered.

Finally, as in WK82, storm strength S_w ² at a given time is defined as the ratio of simulated maximum updraft

² Storm strength is represented by S in WK82, but named S_w herein to avoid confusion with SAL components.

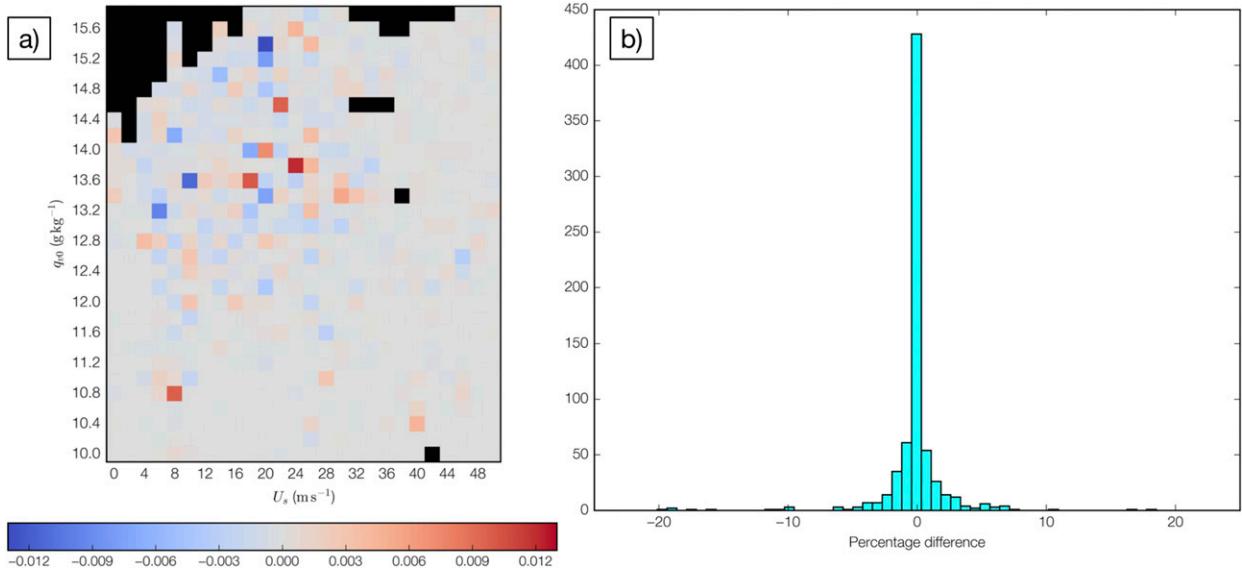


FIG. 3. Differences in 60-min storm strength (also shown in Fig. 11b) between the two microphysics schemes in (a) absolute and (b) relative terms of the interscheme mean for that cell. Black pixels in (a) mark simulations that crashed early in either microphysics scheme. The y axis in (b) indicates histogram-bin count (percentage categories as labeled on the x axis).

speed W_{\max} to the maximum that could be achieved through moist adiabatic ascent (assuming parcel-theory concepts as in WK82):

$$S_w = \frac{W_{\max}}{\sqrt{2\text{CAPE}}}. \quad (7)$$

3. Results: Diagnostic heat maps

In the following sections, heat maps are presented in the manner outlined schematically in Fig. 2, with the σ uncertainty of each field contoured in black (other than Figs. 12 and 15). As each simulation domain is symmetrical along the y -axis center, only one-half of the domain data were considered during statistical processing to reduce the computational expense, as in WK82. The differences in diagnostics stemming from using Morrison or NSSL two-moment microphysics scheme were close to universally negligible. Almost all simulation differences were within 5% of the pair’s mean; for example, differences in 60-min S_w are presented in Fig. 3. In this example, the small number (<10) of outliers do not noticeably change the nature of each scheme’s heat map. This pattern is common with all other diagnostic fields that appear herein. After checking difference covariance matrices and comparing each microphysics’ diagnostic-value distributions (not shown), the author is confident that results solely from the Morrison scheme can be presented and interpreted without loss of applicability to the NSSL

two-moment scheme (or in general). This is further addressed with the slight smoothing of the σ -uncertainty fields, as discussed in section 2.

a. Wind and rotation fields

The q_{v0} - U_s phase space for maximum near-surface (~ 200 m AGL; Fig. 4) wind suggests that wind speeds potentially dangerous to aircraft, structures, etc. ($>25 \text{ m s}^{-1}$; Fujita 1990) occur in high to extreme values of CAPE, as found by Grams et al. (2012). Further, within this subregion of phase space, there are pockets of subsevere near-surface wind associated with $U_s \sim 8$ and $\sim 22\text{--}30 \text{ m s}^{-1}$. This intermittency across phase space is likely the fingerprint of chaos: nonlinear, abrupt changes in sensible weather from small, linear increases in shear. Two local maxima exist in the uncertainty field σ_{wind} , both near the maxima in near-surface wind magnitude (Fig. 4): one at low to midshear, high buoyancy; another at extreme shear, midbuoyancy. In between, there is a col across a wide region of mid- to high buoyancy and shear in phase space. Again, this highlights the variability in predictability across q_{v0} - U_s phase space.

When we inspect the diagnostics for maximum updraft (Fig. 5a) and downdraft (Fig. 5b) speeds, we see a U-shaped lobe of higher ($>15 \text{ m s}^{-1}$) updraft speed at larger q_{v0} and U_s values. In updraft data (Fig. 5a), the two σ -uncertainty maxima at 1) low- to midshear, low- to midbuoyancy, and 2) extreme-shear, low-buoyancy regimes are shifted toward lower-buoyancy phase space, when contrasted with σ_{wind} . These locations change again

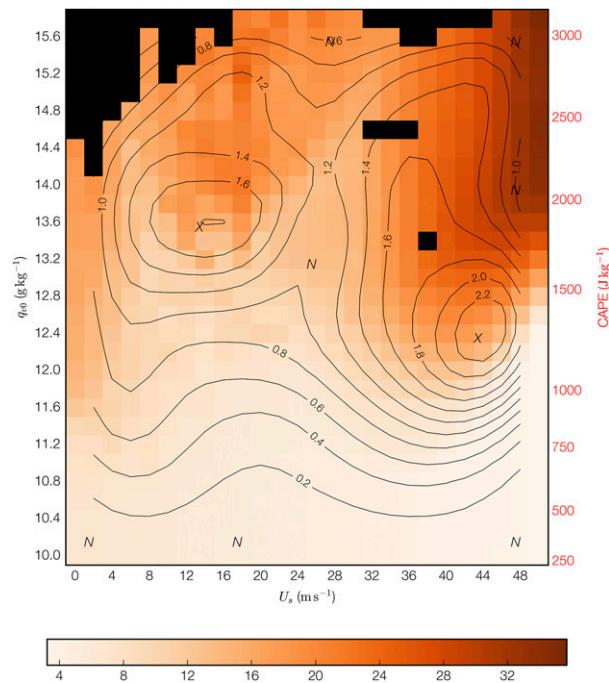


FIG. 4. Heat map in q_{v0} - U_s phase space, where each colored pixel represents the maximum lowest-level wind speed at any point or time for that given simulation, for the Morrison microphysics scheme. The darker the color, the stronger the wind, as denoted by the legend (m s^{-1}). Note the color scheme is normalized for each time's extrema to accentuate areas of largest uncertainty per time. Black lines contour σ uncertainty, which measures sensitivity to small changes in the initial environmental profile, and is contoured every 0.2 m s^{-1} . Local maxima and minima in σ uncertainty are denoted by X and N symbols, respectively.

for downdrafts (Fig. 5b), with σ uncertainty located toward higher buoyancy values. Strong downdrafts are also nearly coincident with some local minima in surface wind maxima, discussed in the previous paragraph (cf. Figs. 4 and 5b at the $U_s = 8$ and 24 m s^{-1} columns and at high buoyancy values).

Next shown in Fig. 6 are heat maps of low- and midlevel vertical relative vorticity (at ~ 925 and 500 hPa , respectively), hereby termed ξ_{925} and ξ_{500} . We again see a U-shaped lobe as in previous heat maps; this is similar to the analogous figure in WK82 (their Fig. 11), as expected from the similarities in experiment configuration. In contrast to near-surface wind and vertical motion diagnostics, we find very little ξ_{925} and ξ_{500} in low-shear environments, confirming the textbook importance that vertical wind shear plays in mesocyclone development (Markowski and Richardson 2010, 201–244). Indeed, at high to extreme levels of shear and buoyancy, high ξ_{925} ($>4 \times 10^{-3} \text{ s}^{-1}$) is associated with these supercellular storms (Fig. 6a). This regime, seemingly ripe for mesocyclonic development, concurs with observational

studies of tornado occurrence (Brooks et al. 2003, their Fig. 1). At midlevels (Fig. 6b), the pattern of ξ_{500} is similar to that in ξ_{925} , but with the phase-space maximum shifted toward lower (but still high) values of U_s . The σ -uncertainty field follows a similar pattern in ξ_{925} (Fig. 6a) compared with ξ_{500} : two prominent σ_ξ maxima in 1) low- to midshear, high-buoyancy and 2) extreme-shear, low-buoyancy regimes; these maxima are both displaced toward higher buoyancy levels at 500 hPa (Fig. 6b).

The maximum in updraft helicity across phase space, as a product of vertical motion and rotation, is likely to occur near the intersection of maxima in Figs. 5 and 6, where the author assumes that strong updrafts collocate with the mesocyclone. This relationship is found in Fig. 7b, where the largest UH_{02} values partly overlap with the largest in ξ_{925} . Accordingly, UH_{25} is more confined to extreme values of both q_{v0} and U_s (Fig. 7a). In both low- and midlevel layers, the uncertainty σ_{UH} is largest 1) on the perimeter of maximum UH_{02} and UH_{25} values in high to extreme shear and high to extreme buoyancy, and 2) in a zone of near-midshear values and near-extreme-buoyancy values. As UH_{02} is frequently used as a signal for tornadogenesis in quasi-operational ensemble forecast systems (e.g., Wheatley et al. 2015; Skinner et al. 2018), these results emphasize how low predictability necessitates adequate ensemble dispersion and reliability, especially in contexts where deterministic thinking is historically preferred (Novak et al. 2008).

b. Precipitation

Figure 8 presents three precipitation-related products: First (Fig. 8a), cubic rainfall accumulation integrated over the entire half domain is presented. As no simulated storm exits their domain boundaries within the 2 h, this product can be interpreted as total precipitation from the thunderstorm in this time. As such, it is unsurprising that Fig. 8a displays a similar U-shaped pattern in both rainfall and contours of uncertainty σ_{rain} seen in the previous storm-related diagnostic heat maps. The broad midshear, high-buoyancy regime is associated with a local minimum in σ_{rain} , implying these substantial storm rainfall totals are relatively predictable in contrast with nearby states in phase space. However, when we consider maximum 2-h accumulated rainfall at any point (Fig. 8b), the pattern changes: there is a bullseye in both magnitude and σ_{rain} at low shear and high to extreme buoyancy. These high pointwise rainfall accumulations are observed here in tandem with high rainfall rate, slow storm speed, and extended storm life³

³ Storm lifetime is not shown explicitly but can be inferred in many plots by the presence of a nonzero diagnostic value.

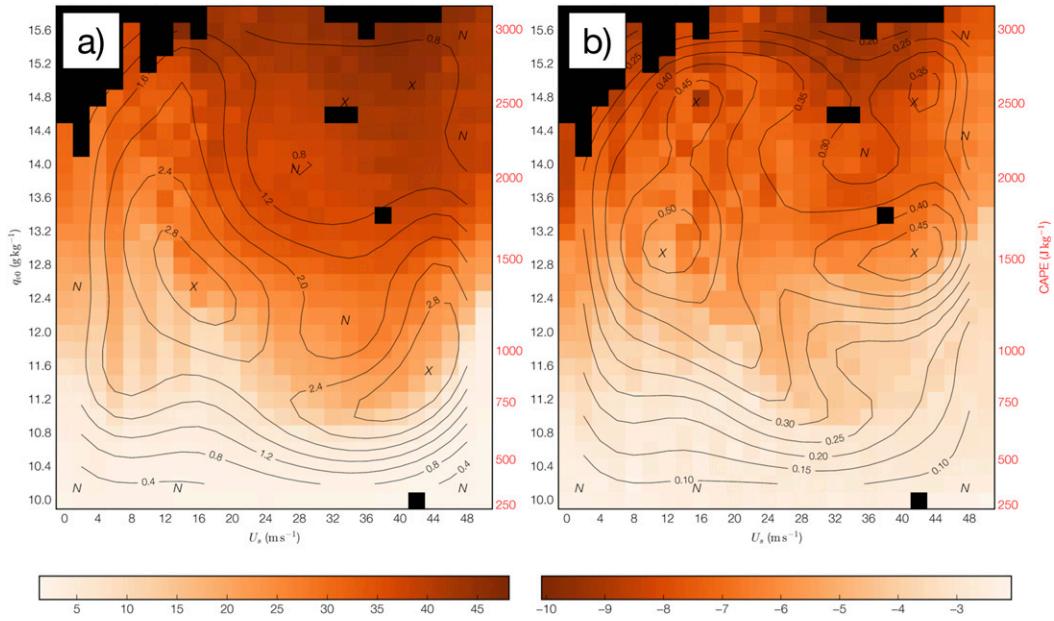


FIG. 5. As in Fig. 4, but for (a) maximum updraft and (b) maximum downdraft (m s^{-1}) at any level, point, or time. The σ -uncertainty field is contoured every 0.4 m s^{-1} in (a) and 0.05 m s^{-1} in (b).

(among other factors; Doswell et al. 1996), evidenced in the following heat maps. There is a weaker σ_{rain} maximum in midbuoyancy, extreme-shear environments (Fig. 8b), approximated collocated with the σ -uncertainty maximum seen in some previous heat maps (cf. Fig. 4a).

Last, maximum 2.5-min rainfall accumulation at any point and time—a proxy here for the storm’s instantaneous

rainfall rate—shows in Fig. 8c that a far broader area of phase space supports heavy rainfall rates $>2 \text{ mm}$ per time step (every 2.5 min; or as an hourly rate, 48 mm h^{-1} or 1.89 in. h^{-1}). The heaviest rates are in low- to mid-shear and high- to extreme-buoyancy environments, corresponding to the aforementioned pointwise accumulation maximum. This area of phase space is also

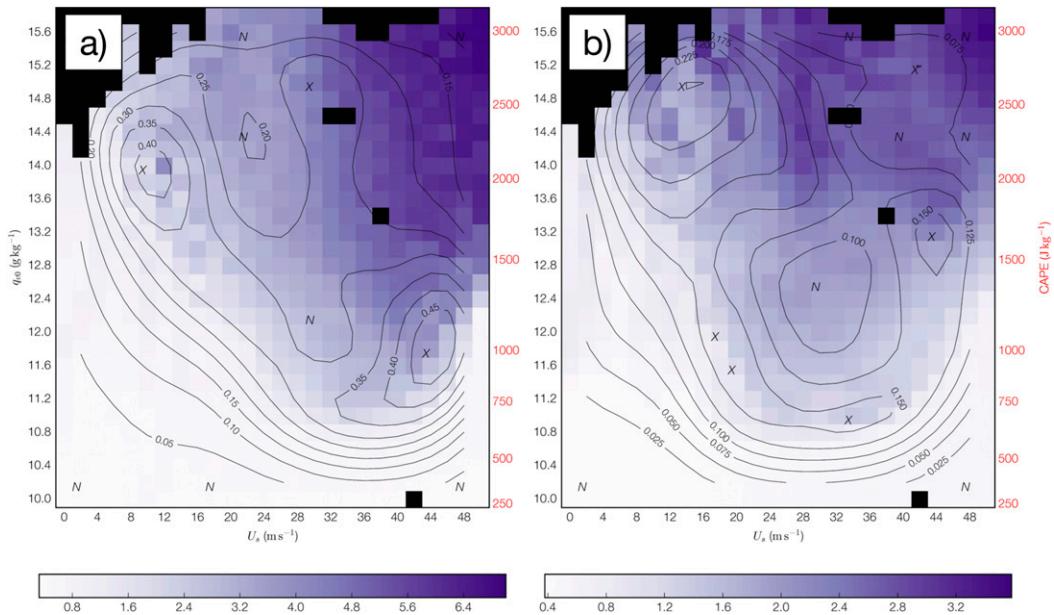


FIG. 6. As in Fig. 4, but for maximum relative vertical vorticity ($\times 10^{-3} \text{ s}^{-1}$) at (a) ~ 925 and (b) ~ 500 hPa for any point or time. The σ -uncertainty field is contoured every $0.05 \times 10^{-3} \text{ s}^{-1}$ in (a) and $0.025 \times 10^{-3} \text{ s}^{-1}$ in (b).

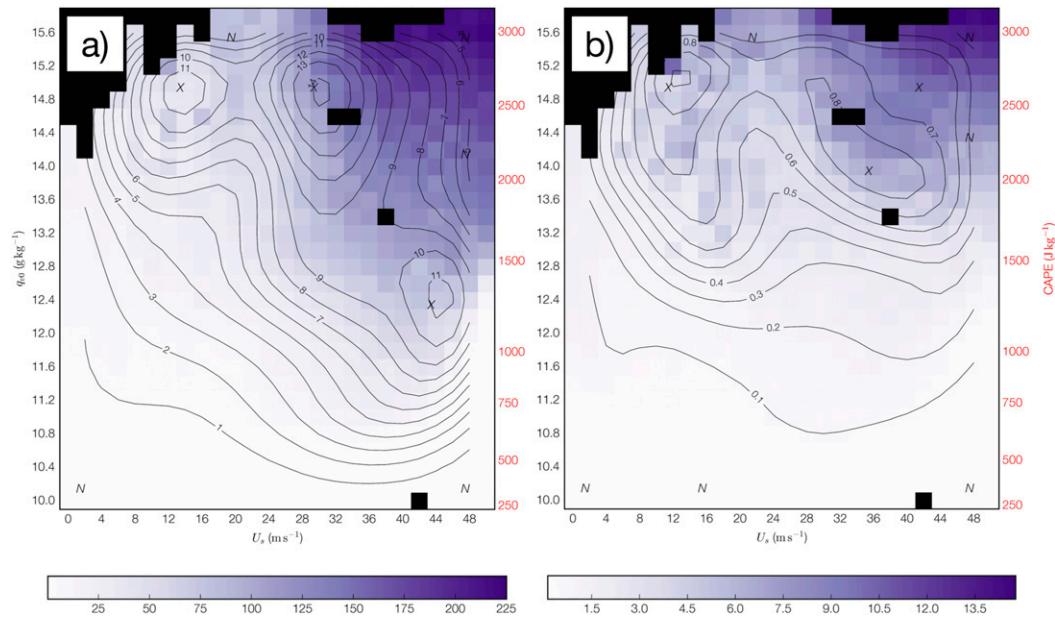


FIG. 7. As in Fig. 4, but for maximum updraft helicity ($\text{m}^2 \text{s}^{-2}$) for any point or time in the (a) 2–5- and (b) 0–2-km layers. The σ -uncertainty field is contoured every $1 \text{ m}^2 \text{ s}^{-2}$ in (a) and $0.1 \times 10^{-3} \text{ m}^2 \text{ s}^{-2}$ in (b).

associated with a maximum in σ_{rain} , which supports the relationship that rare or extreme weather events are also the most unpredictable (Sterk et al. 2016). Finally, the low-shear, low-buoyancy environments for all three rainfall products (Fig. 8) support substantially smaller rainfall accumulations and rates. The following discussion details further storm properties that reveal a bifurcation in phase space: between storms that begin to dissipate at 30–60 min, and those that continue to strengthen over the 2 h.

c. Storm speed, size, strength

Figure 9 presents the speed of detected convective cells at 30, 60, and 90 min of simulation time. The detection methodology is discussed in section 2, and instances where a storm was not detected are represented by a black pixel in Fig. 9. After 30 min (Fig. 9a), only low to midshear and high to extreme CAPE supports detectable cell development. This echoes WK82 (their Fig. 3), where storms decrease in updraft strength as U_s increases, at 30 min of

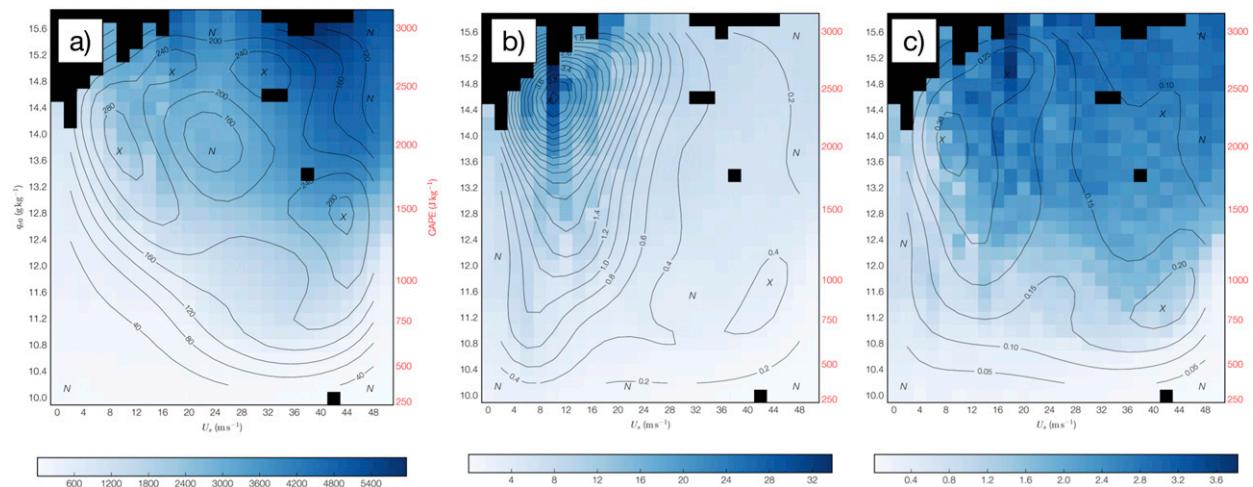


FIG. 8. As in Fig. 4, but for (a) rainfall integrated over the entire half domain (10^3 m^3), (b) maximum 2-h accumulated rainfall at any point (mm), and (c) maximum 2.5-min accumulation for any point and time (mm). The σ -uncertainty field is contoured every $40 \times 10^3 \text{ m}^3$ in (a), 0.2 mm in (b), and 0.05 mm in (c).

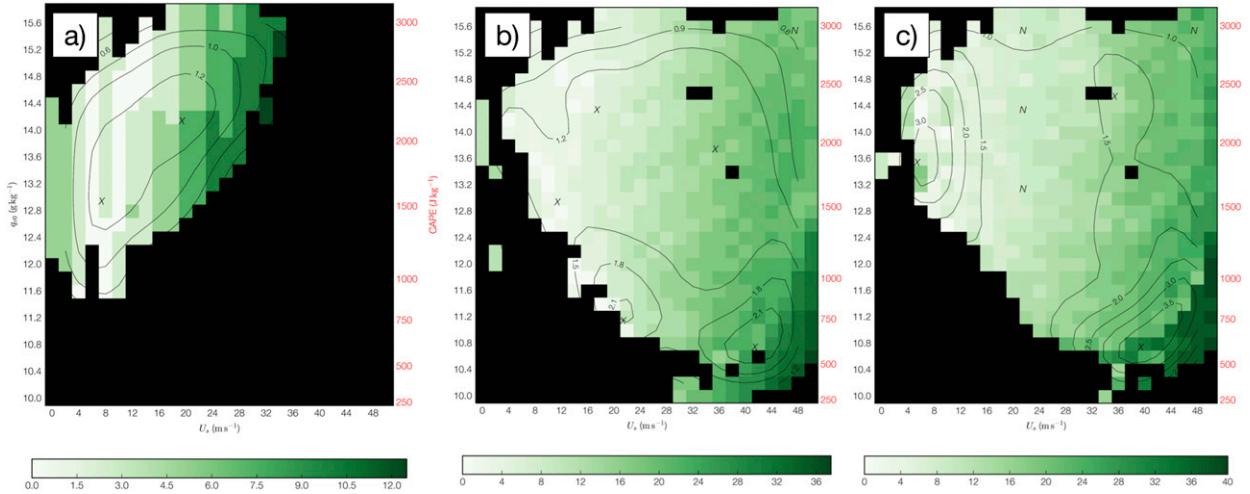


FIG. 9. As in Fig. 4, but for thunderstorm speed (m s^{-1}) at (a) 30, (b) 60, and (c) 90 min. The σ -uncertainty field is contoured every 0.2 m s^{-1} in (a), 0.3 m s^{-1} in (b), and 0.5 m s^{-1} in (c). Thunderstorms were detected if the location of the maximum 500-hPa updraft at a given time corresponded to simulated composite reflectivity greater than 25 dBZ; runs with no detected storms are represented here with black pixels.

simulation. Uncertainty of storm speed is highest in the center of this regime, as marked by the local maximum in uncertainty σ_{speed} , and speeds increase from almost stationary to over 10 m s^{-1} (19.4 kt) as shear increases. The computed storm speed is tightly related to U_s , as in observed systems (Bunkers et al. 2000). After 60 min (Fig. 9b), storms are detected in most regimes. Storm speed increases fairly linearly with U_s , though the highest σ_{speed} is found in low to midbuoyancy at extreme U_s values, associated with the highest speeds ($>30 \text{ m s}^{-1}$, or 58.3 kt). This

is similar to Kirkpatrick et al. (2007), where thunderstorms developing within this area of phase space moved at speeds approaching the environmental 0–6-km bulk vertical wind shear, unlike relatively slower cells in regimes possessing high shear and buoyancy. By 90 min (Fig. 9c), the pattern is similar to 30 min earlier (in Fig. 9b), though 1) there is a much stronger gradient between fast (25 m s^{-1}) and faster (35 m s^{-1}) cells at extreme-shear values and low buoyancy and 2) σ_{speed} has increased in the high-buoyancy, low-shear regime. This second point suggests that location errors

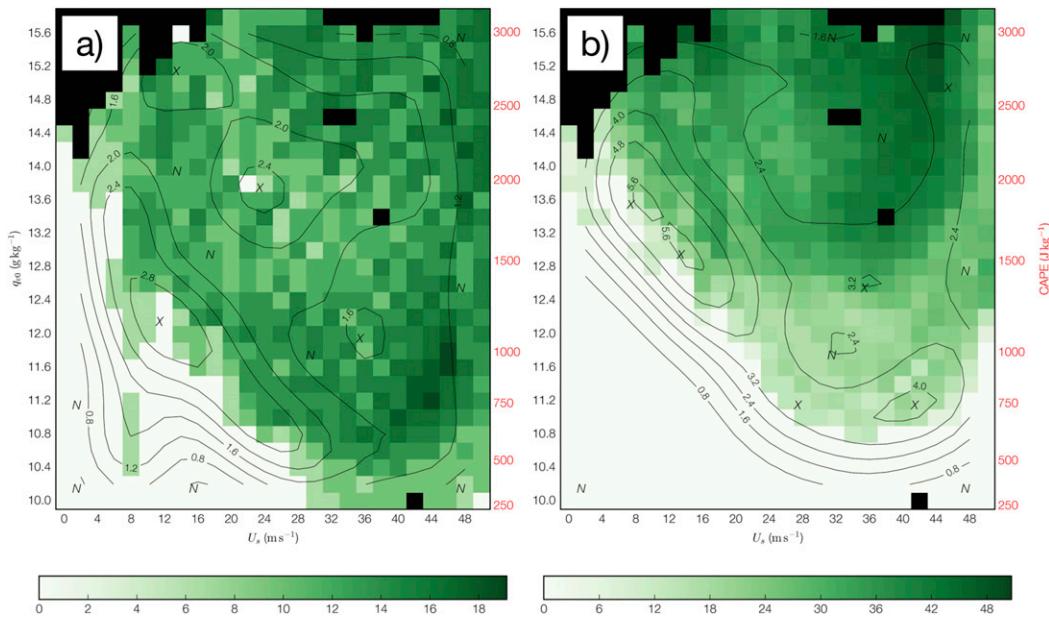


FIG. 10. As in Fig. 4, but for equivalent thunderstorm diameter (km) at (a) 60 and (b) 120 min. The σ -uncertainty field is here contoured every 0.4 km in (a) and 0.8 km in (b).

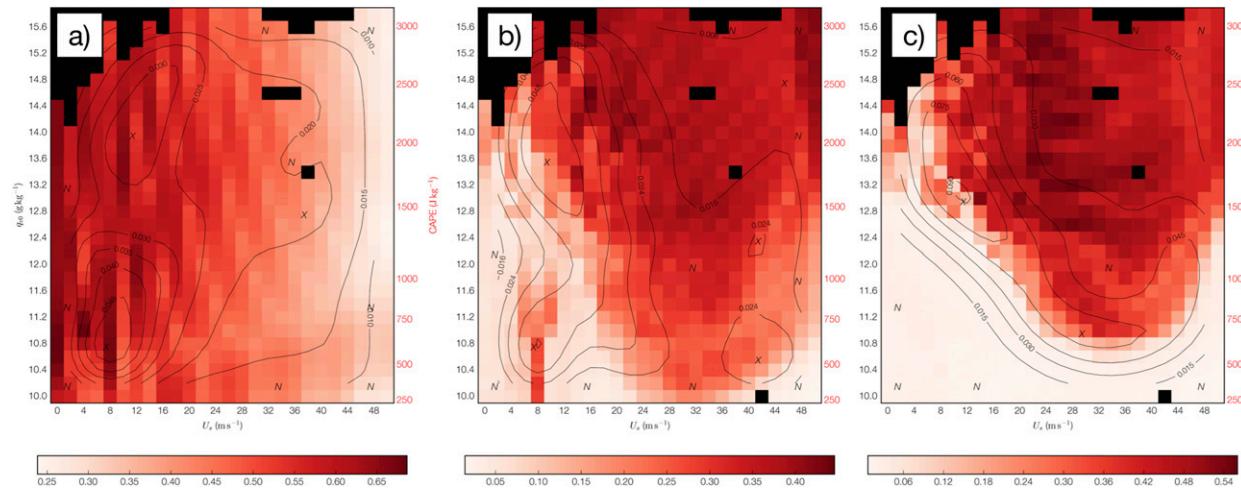


FIG. 11. As in Fig. 4, but for storm strength (unitless) at (a) 30, (b) 60, and (c) 90 min. The σ -uncertainty field is here contoured every 0.005 in (a), 0.008 in (b), and 0.015 in (c).

may become especially prominent in the latter stages of very short-term (0–6 h) hazardous weather forecasts in operations-ready ensemble systems (Yussouf et al. 2016; Lawson et al. 2018).

Equivalent storm diameter is shown in Fig. 10 (the formulation is discussed in section 2). This heat map shows much uncertainty in storm size at 1 h (Fig. 10a), with storms detected⁴ at almost all regimes away from the lowest q_{v0} and U_s values. The considerable variation between adjacent pixels likely stems from phase errors in storm growth (not shown). As time progresses (at 2 h; Fig. 10b), larger thunderstorms are more confined to higher q_{v0} and U_s values bounded by the U-shaped perimeter. There is a local minimum of σ uncertainty in mid-shear, mid-buoyancy environments, suggesting storms in this regime grow in size fairly similarly, regardless of small IC error. The sharpest gradient of storm size (maximum σ_{size}) occurs on the low-shear side of the U-shaped perimeter, which is near the aforementioned zones of high magnitude and uncertainty in rainfall and storm-speed products (cf. Figs. 8–10).

To summarize numerous diagnostic concepts in one product, storm strength S_w is computed as in WK82 [repeated here in Eq. (7)], and results for three lead times are presented in Fig. 11. As S_w is normalized by the square of twice CAPE, a value of unity represents a storm that is as strong as the profile of moist ascent can support (from the tenets of parcel theory); the smaller the fraction, the weaker the storm, but normalized by its

location in phase space. After 30 min (Fig. 11a), we find that lower-shear regimes support much quicker storm evolution. Taken as a snapshot in time, note the large variation in adjacent pixels for all low to high values of U_s . This may be associated with the lower skill measured in operational forecasts within the same regime (Herman et al. 2018). At 60 min (Fig. 11b), there has been a considerable collapse in S_w values in the low-shear, low-buoyancy regime: for instance, S_w ratios fall from 0.5 to 0.65 to below 0.3 in the 30 min between Figs. 11a and 11b. We see a familiar U-shaped perimeter dividing storms that become stronger as time progresses (to the high-shear, high-buoyancy side of phase space) and those that dissipate over the 30–60-min time period (in the low-shear, low-buoyancy regime; white or light pixels in Fig. 11b). By 90 min (Fig. 11c), there is a sharp divide between the two basins: this U-shaped perimeter can be interpreted as a tipping point. While the largest variation in storm strength σ_{S_w} is found in the low-shear, high-buoyancy regime (cf. previous discussion on precipitation), the σ_{S_w} maximum extends around the perimeter into the high-shear, low-CAPE regime (e.g., Sherburn and Parker 2014; Mulder and Schultz 2015; Sherburn et al. 2016). This corroborates Herman et al. (2018), who found forecasts within the high-shear, low-buoyancy regime were poor. These results confirm that the sensitivity is acute when moving from this regime to a slightly more unstable profile: a matter of $O(100)$ J kg^{-1} discriminates between strengthening and dissipating cells. A similarly high sensitivity of storm evolution to small errors was also reported by Crook (1996) and Elmore et al. (2002), who found that critical differences in initial-condition and representivity errors were on the order of observation errors. This bifurcation

⁴ Profiles with an equivalent diameter less than ~ 9 km are below the effective resolution of the simulation, and caution must be taken when interpreting this diagnostic as a result.

(solutions splitting into separate watersheds) in 90-min storm strength will be shown further through time series of updraft strength in the next section.

The discrepancy between profiles with no storm detected in Fig. 9 and those with nonzero storm attributes (e.g., low-buoyancy, low-shear regimes in Figs. 11a,b) is caused by the reflectivity threshold in the storm-detection methodology, during storm-speed estimation. At the lowest U_s and q_{v0} values, updrafts still occur, but storms are either poorly resolved at 3-km grid spacing, or are so weak that they do not pass the threshold. It is unlikely that cellular convection in the bottom left of each heat map would be considered a thunderstorm at any point in the simulation (no lightning parameterization is used).

When animated for every output time (not shown), the high variability in low-shear regimes is represented by local maximum–minimum couplets of S_w rippling through phase space. While caution must be taken when interpreting results at this horizontal grid spacing, this may be a signal of a pulsating or cyclical behavior, also seen in time series of vertical motion (presented in the next section).

4. Predictability synthesis

Contours of σ uncertainty have been shown in the previous section to highlight potential tipping points in phase space. We now turn to two mathematical quantities to determine regions of phase space with low predictability in both the gridpoint and object-based senses.

a. Object-based predictability (SAL)

The following heat maps show standard deviation of each SAL component generated from the simulated composite reflectivity fields. This quantity represents the uncertainty in the nature of object-based differences (see section 2). Again, there is little difference in SAL-component character for both Morrison and NSSL two-moment matrices. Both are presented in Fig. 12 for comparison, but as the differences are negligible, the following discussion arbitrarily addresses the Morrison microphysics scheme.

The structure (purple; Figs. 12a,d) and amplitude (green; Figs. 12b,e) components trace out U-shaped maxima, resembling the perimeter of the U-shaped lobe in previous heat-map figures. Away from this U-shaped maximum and the low-shear, low-buoyancy maximum in the bottom left, there is little variation in object structure S and domain-integrated reflectivity A . These results suggest that thunderstorm structure and strength, in an object-based sense, are less predictable along the U-shaped perimeter. In contrast, the standard deviation of location component L yields little systematic pattern, though highest variation occurs in the high-shear,

mid- to high-buoyancy regime. The lack of U-shaped pattern in L , in contrast to S and A components, is related to the asystematic errors in storm speed (Fig. 9), and suggests that the predictability “field” of, say, storm speed is very different from that associated with other (wind, rainfall, etc.) diagnostics.

As seen in Fig. 11, the U-shaped perimeter appears to discriminate between weakening and strengthening cellular convection. As in WK82, a wide variety of storm strengths can be seen across phase space. To demonstrate whether this U-shape acts as a tipping point, two contrasting thunderstorm life cycles are shown in Fig. 13, taken from two runs adjacent in phase space, and whose only difference in profile is 0.2 g kg^{-1} ($\sim 125 \text{ J kg}^{-1}$ CAPE equivalent). The two runs lie on the perimeter of the U-shaped maximum, marked by the two left-side pixels within the red box in Fig. 12. Differences between the supercells’ appearance in simulated composite reflectivity become apparent between 80- and 100-min simulation time (Fig. 13), and by 120 min, continued strengthening is particularly evident in the simulation with slightly higher buoyancy (Fig. 13b).

Similarly, Fig. 14 shows maximum vertical motion at model level 13 (around 500 hPa or 5.5 km AGL), for the same two simulations and two additional adjacent simulations that straddle the U-shaped maximum in Fig. 12 (again, at the location of the red box). The time series show the initial explosive growth of the storm cell around 20 min, a gradual decrease in intensity between 40 and 60 min, and then a bifurcation of solutions: the two simulations with slightly larger buoyancy generate pulsating and strengthening updrafts with a period of around 20 min, whereas the slightly lower buoyancy results in a dissipating of the updrafts with somewhat longer (30 min) period of pulsation.

Keeping wind shear constant, and increasing buoyancy (i.e., moving in the positive y direction), the wavelength and amplitude of the oscillations decrease after 60 min (not shown). Keeping buoyancy constant and increasing or decreasing shear dampens the oscillation and chokes the updraft, moving into the basin of dissipation within phase space. Hence, the U-shaped maxima in S and A components of SAL (Fig. 12) delineate the two basins of attraction in buoyancy–shear space. Away from this perimeter, storm attributes diverge quickly, and very different 2-h forecasts of similar atmospheric states ensue: this is the fingerprint of chaotic growth.

b. Gridpoint predictability (DKE)

In contrast to SAL, DKE measures gridpoint-scale predictability destruction, with no tolerance of small storm-location differences. The unit–time derivative of

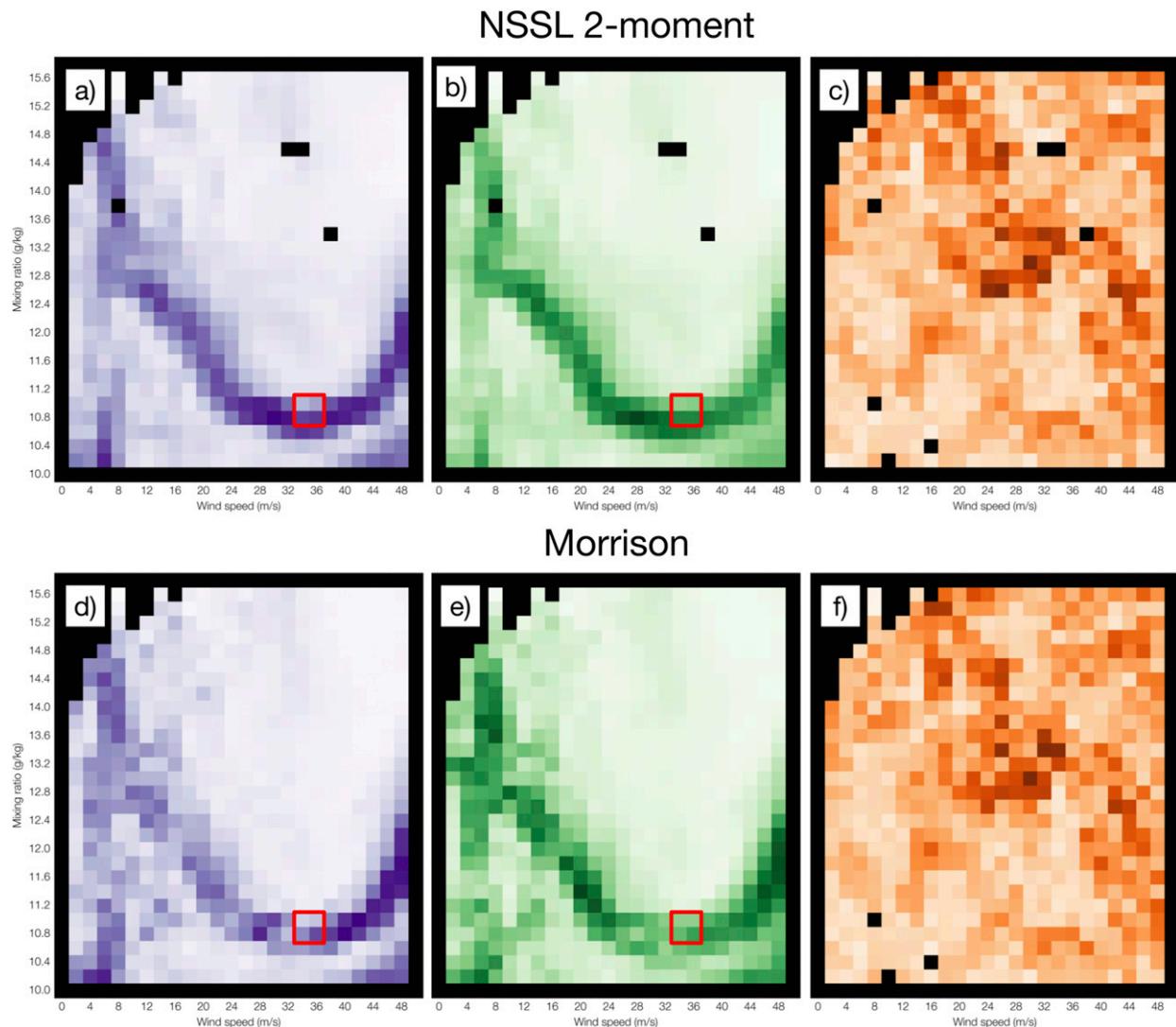


FIG. 12. Heat maps showing uncertainty in SAL components at 2 h of simulation time for the (a)–(c) NSSL two-moment and (d)–(f) Morrison microphysics schemes. Panels show the standard deviation of (a),(d) structure, (b),(e) amplitude, and (c),(f) location components (unitless; the values are functions of domain size and not meaningful here, hence neglected). Red boxes indicate the four cells that are shown in the time series of maximum 500-hPa vertical motion in Fig. 14.

uncertainty (dDKE) highlights regions in which predictability is being destroyed the fastest. Values of dDKE from the Morrison scheme are presented in Fig. 15. In the first 30 min (Fig. 15a), there is rapid DKE growth in the high-buoyancy, low-shear regime. This appears as a rippling pattern (when animated; not shown) that evokes the pattern observed in the storm strength diagnostic (Fig. 11). By 60 min (Fig. 15b), the highest dDKE is occurring in the high-buoyancy, high-shear regime, associated with the long-lasting supercells occurring here. As seen in the animation (not shown), waves of “predictability destruction” emanate from the area of maximum dDKE, most clearly at 90 min (Fig. 15c). These waves may represent the phase

difference between adjacent cells in supercell growth and mesocyclonic cycling with different cycling periods (Burgess et al. 1982; Adlerman et al. 1999). Signals of this pulsation can be seen as sinusoidal waves in 500-hPa vertical motion (Fig. 14). However, the 20–30-min periods observed herein are somewhat shorter than those simulated in Adlerman and Drogemeier (2002, their Fig. 8). From their conclusions on grid spacing, it is also surprising that a 3-km grid could resolve such cycling, though recent work supports such a claim (Britt et al. 2018).

The simulations that support strong supercells in the extreme regimes of shear and buoyancy dominate dDKE magnitude; as a result, the DKE (not shown) and

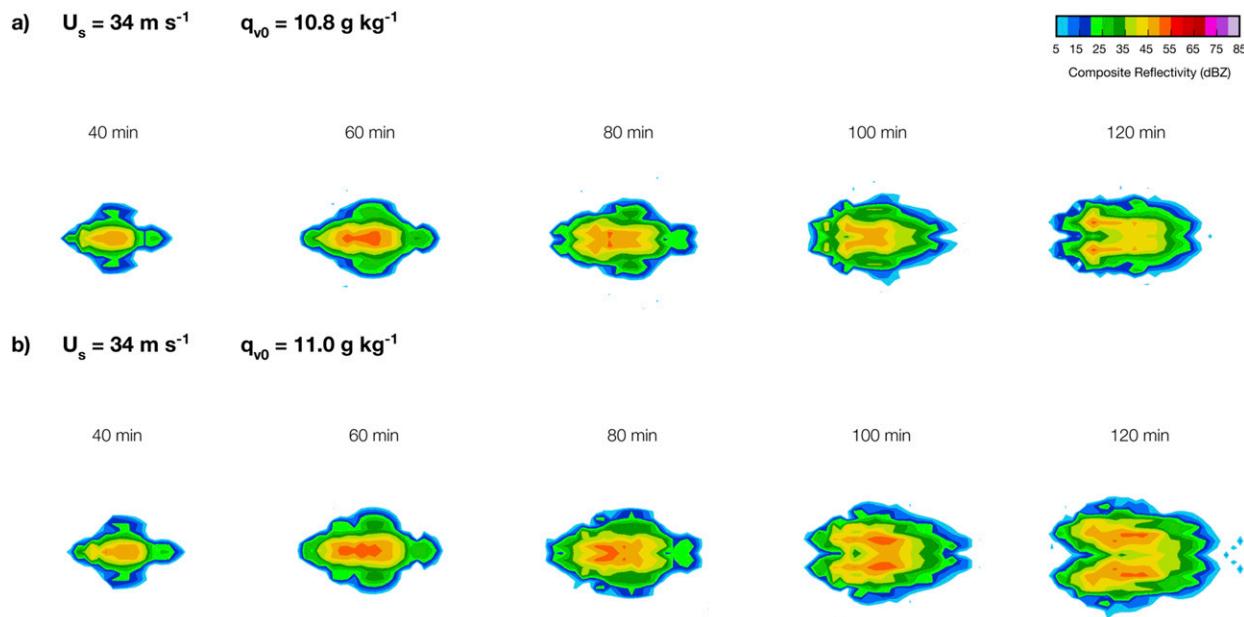


FIG. 13. Life cycle of two simulated thunderstorms shown every 20 min in simulated composite reflectivity for the (a) $U_s = 34 \text{ m s}^{-1}$, $q_{v0} = 10.8 \text{ g kg}^{-1}$ and (b) $U_s = 34 \text{ m s}^{-1}$, $q_{v0} = 11.0 \text{ g kg}^{-1}$ simulations using Morrison microphysics (whose location in phase space is shown by the red box in Fig. 12). The 0- and 20-min panels are omitted because of simulated reflectivity less than 0 dBZ. Scales for both panels are identical; thunderstorm location is centered within each frame.

dDKE (Fig. 15) fields look similar beyond around 90 min (not shown). Given dDKE is a proxy for predictability loss, we find that the *strongest storms are associated with the fastest traditional (gridpoint) predictability destruction*. In some ways, this is expected: if the bifurcation were between no storms at all and supercells, then an ensemble forecast ran in an environment devoid of convection would be very predictable (low ensemble spread) compared to those with strong thunderstorms (high ensemble spread). However, it is generally a counterintuitive result, given that supercells are a more persistent phenomenon than, for example, single cells, and long suspected to contain mechanisms that suppress (object-based) predictability destruction (Lilly 1986). As the storm cycles and strengthens, gridpoint predictability is more efficiently destroyed through chaotic error growth. However, the object-based predictability may be extended in the manner discussed by Lilly (1986): a strong rotating supercell insulates the storm from simply detraining and dissipating. In other words: the thunderstorm’s mode and characteristics are more predictable than its gridpoint evaluation (with zero tolerance.) While Lilly’s ideas relating to helicity and suppression of chaotic error growth was later shown to be overly optimistic (Lilly 1990), his suspicion of extended supercell predictability was confirmed by Drogemeier and Levit (1993), and the mechanism that yields a longer object-based

predictability estimate may be related to nonlinear feedback loops (Shen et al. 2018), or continuous propagation on the storm’s flank as a consequence of midlevel mesocyclogenesis and a strengthening low-level updraft [see discussion in Weisman and Rotunno (2000, p. 1454)].

The phase error that may be associated with the waves in dDKE would result in unfair punishment from not only traditional point-to-point verification metrics, but also neighborhood schemes that include windows of location but not time [a review and critique of neighborhood schemes can be found in Schwartz and Sobash (2017)]. Scores such as the extended fractions skill score, published independently by Schwartz et al. (2010) and Duc et al. (2013), apply the windowing over a number of adjacent time steps (as well as considering the ensemble member dimension), and are recommended. It follows that one could subset a climatology such that it were confined to that region of phase space, and use this to more appropriately estimate the upper bound of predictability. In the case of quasi-operational and experimental forecast systems, however (e.g., Wheatley et al. 2015), it begs the question of defining a model climatology with an insufficient sample of archived cases.

5. Conclusions

This study has presented theoretical predictability for idealized cellular convection across a high-granularity

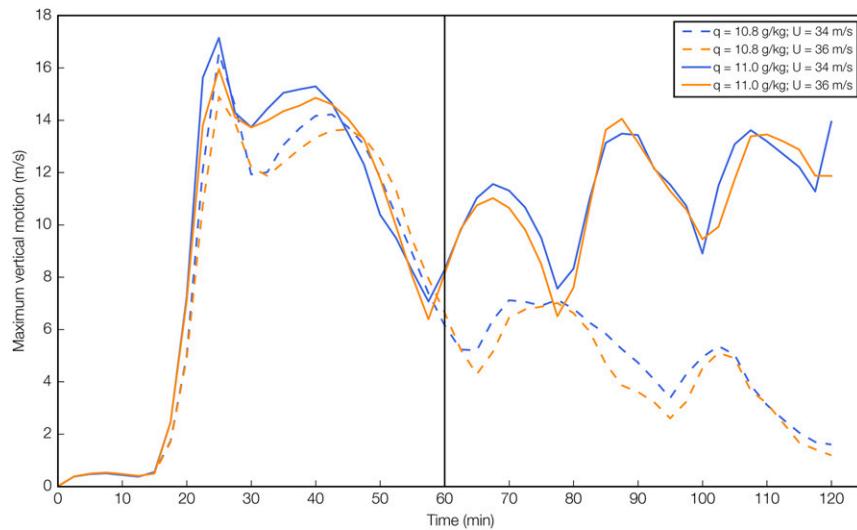


FIG. 14. Time series of maximum vertical motion at model level 13 (~ 500 hPa or ~ 5 km AGL) for the Morrison microphysics scheme. These four simulations, whose location is marked with a red box in Fig. 12, show a bifurcation between strengthening and dissipating supercells occurring around an hour after initialization. The black vertical line indicates an hour of simulation time. The dashed lines represent storms embedded in atmospheres containing $\sim 100 \text{ J kg}^{-1}$ less CAPE than those denoted by solid lines (which tend to dissipate after 1 h). Also, blue lines indicate storms embedded in flow with a 2 m s^{-1} weaker 0–6-km wind difference.

phase space of vertical wind shear and buoyancy. After an hour of simulation time, storms can be split into two groups, separated by a U-shaped bifurcation in phase space (Fig. 16). This bifurcation suggests two basins of attraction: one basin at higher buoyancy values, associated with storms that continue strengthening, and another associated with dissipating convection unsupported by the environment, at lower buoyancy values. While it is expected that more CAPE and shear available to a storm

will usually increase its strength, this increase does not occur gradually, but abruptly around the tipping point. This suggests ensemble forecasts of isolated thunderstorms in the low- to midbuoyancy regime—for all shear values—may be associated with large uncertainty (i.e., spread) as an inherent character of the regime predictability. Two areas of phase space were frequently sensitive to small IC changes: 1) the low-buoyancy, high-shear regime and 2) low-shear regime at all buoyancies.

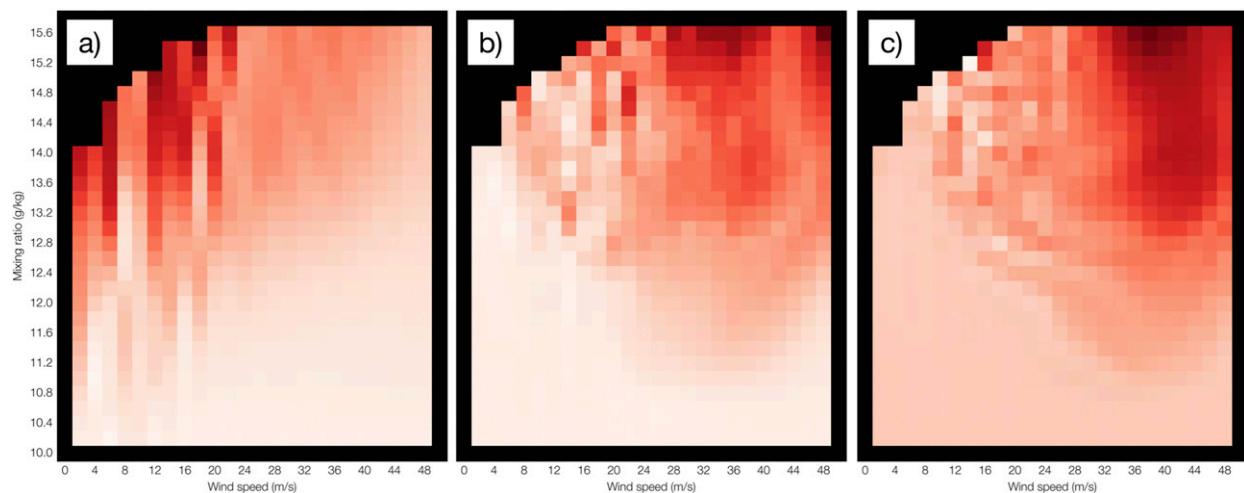


FIG. 15. As in Fig. 5, but for $d\text{DKE}$ (the time derivative of DKE) for the Morrison microphysics scheme at (a) 30, (b) 60, and (c) 90 min of simulation time. Large $d\text{DKE}$ values represent higher rates of gridpoint predictability destruction. Magnitudes of $d\text{DKE}$ are not meaningful here, and neglected for clarity.

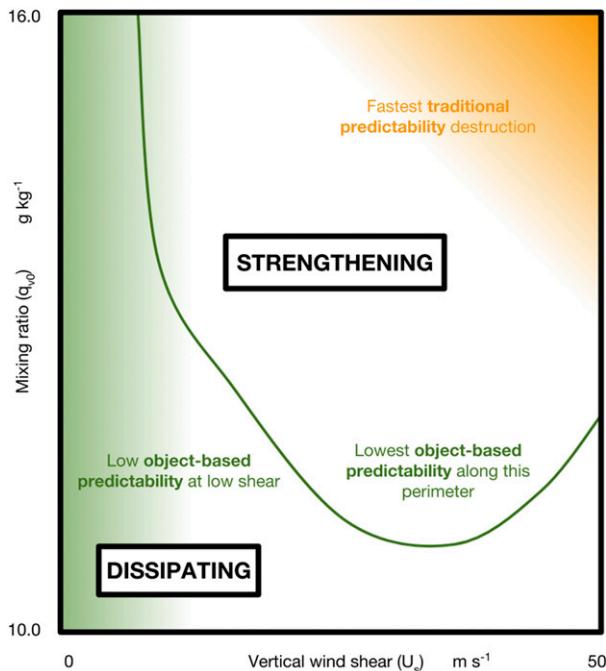


FIG. 16. Schematic of general results detailed herein. The green curve denotes the U-shaped maximum (e.g., Fig. 12a) that marks the boundary between the strengthening and dissipating solutions (black text). The region of phase space that follows this green curve (and the swath of green for all low-shear values) is associated with lower object-based predictability. In contrast, the orange shading denotes the region of phase space with the maximum $dDKE$, or traditional predictability destruction, which decreases as one moves away from the top-right corner.

Both were identified as regimes with less skillful forecasts made by the U.S. Storms Prediction Center (Herman et al. 2018).

To further assess theoretical predictability, a distinction between two senses of predictability destruction was made: (i) one that is measured by DKE growth, and maximized in the largest storms at extreme buoyancy and shear levels; and (ii) another measured by object-based approaches, maximized near tipping points at the discriminator in phase space between strengthening and dissipating storms. The former sense is associated with traditional gridpoint-to-gridpoint forecast error (gridpoint predictability), whereas the latter is associated with error between cellular features, with tolerance for small phase errors (object-based predictability). The latter error is more intuitive when discussing convective-scale forecasts, and tallies with tolerances considered by human forecasters (i.e., when storm mode is more important than a precise location). As such, practical predictability estimates at the convective scale should mimic the forecaster’s mental filtering processes, lest they profoundly underestimate

the utility remnant even after gridpoint error saturation has been reached.

Other findings include the following:

- There is very little difference between the two double-moment microphysics schemes, which suggests that (i) results are likely applicable to other model configurations, and given that microphysics error is a primary source of model error, (ii) predictability is primarily a function of IC error at thunderstorm time scales.
- A rippling pattern in phase space, seen in gridpoint-predictability destruction, may be interpreted as phase offsets of updraft oscillations whose periods change with shear and buoyancy.
- As revealed by an object-based decomposition, the structure and amplitude (or strength) of the thunderstorm is profoundly sensitive to shear and buoyancy in regimes of phase space near the U-shaped perimeter; conversely, there is little pattern to storm-location or storm-speed uncertainty in phase space, suggesting that a strong tipping point may not exist for temporal and spatial phase errors in storm forecasts.
- The maximum rainfall accumulation occurs at low shear and high to extreme CAPE values, whereas total storm rainfall is highest in extreme buoyancy and shear regimes.
- Low-level updraft helicity, a proxy for tornadogenesis, is maximized at the highest U_s and q_{v0} .

In summary, these results should continue to motivate the use of ensemble forecasting systems at the convective scale. It is paramount that these ensemble systems are reliable in order to fully explore the phase space of extreme and rare events; as such, system designers should consider methods that further account for uncertainty, such as stochastic perturbations (Berner et al. 2017), probabilistic postprocessing (e.g., Bröcker and Smith 2008), and so on.

There are inherent shortcomings of this experiment, not least that idealized studies are not immediately applicable to the observed world. Results above must be interpreted within the scope of the experimental configuration: the prescription of a straight hodograph, the limitation of 3-km horizontal grid spacing, and a fixed parameterization suite (bar microphysics diversity). The imposition of a warm bubble also precludes any statement on convective initiation, cell mergers, and so on; more strongly forced phenomena, such as quasi-linear convective systems, may be partly sheltered from the IC sensitivity presented herein. Conversely, horizontally homogeneous lateral-boundary conditions do not introduce any strong forcing (upper-level troughs) events that may improve predictability in real-world dynamics (Anthes et al. 1985). Furthermore, model error is

neglected here other than microphysics diversity. These caveats aside, the signal appears genuine: results herein corroborate observational climatologies, forecast-center verification studies, and the coarser matrices of phase space mapped by WK82.

Future work may repeat the above procedure for a range of sounding types, such as in Kirkpatrick et al. (2007, 2011), or investigate a link between the idealized literature with operational datasets from severe weather forecasting centers.

Acknowledgments. The author thanks two anonymous reviewers, Cody Kirkpatrick, and Robert Fovell for valuable contributions to the manuscript's clarity and content; Ted Mansell and Harold Brooks for improving earlier versions of the manuscript; Corey Potvin, Timothy DelSole, Louis Wicker, Jeffrey Duda, Jack Kain, Nusrat Yussouf, and William Gallus for useful conversation; and Gerald Creager and Oklahoma University OSCER for computational assistance. The WRF runs contained herein were created, submitted, and data mined by open-source Python packages. These include numpy, matplotlib, scipy, netcdf-python, and the author's open-source ensemble verification and creation (evac) codebase found online (<https://github.com/johnrobertlawson/evac>). Funding was provided by NOAA/Office of Oceanic and Atmospheric Research under NOAA–University of Oklahoma Cooperative Agreement NA11OAR4320072, U.S. Department of Commerce.

REFERENCES

- Adlerman, E. J., and K. K. Droegemeier, 2002: The sensitivity of numerically simulated cyclic mesocyclogenesis to variations in model physical and computational parameters. *Mon. Wea. Rev.*, **130**, 2671–2691, [https://doi.org/10.1175/1520-0493\(2002\)130<2671:TSONSC>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<2671:TSONSC>2.0.CO;2).
- , —, and R. Davies-Jones, 1999: A numerical simulation of cyclic mesocyclogenesis. *J. Atmos. Sci.*, **56**, 2045–2069, [https://doi.org/10.1175/1520-0469\(1999\)056<2045:ANSOCM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1999)056<2045:ANSOCM>2.0.CO;2).
- Ancell, B., and G. J. Hakim, 2007: Comparing adjoint- and ensemble-sensitivity analysis with applications to observation targeting. *Mon. Wea. Rev.*, **135**, 4117–4134, <https://doi.org/10.1175/2007MWR1904.1>.
- Anthes, R. A., Y.-H. Kuo, D. P. Baumhefner, R. M. Errico, and T. W. Bettge, 1985: Predictability of mesoscale atmospheric motions. *Advances in Geophysics*, Vol. 28B, Academic Press, 159–202, [https://doi.org/10.1016/S0065-2687\(08\)60188-0](https://doi.org/10.1016/S0065-2687(08)60188-0).
- Berner, J., T. Jung, and T. N. Palmer, 2012: Systematic model error: The impact of increased horizontal resolution versus improved stochastic and deterministic parameterizations. *J. Climate*, **25**, 4946–4962, <https://doi.org/10.1175/JCLI-D-11-00297.1>.
- , and Coauthors, 2017: Stochastic parameterization: Toward a new view of weather and climate models. *Bull. Amer. Meteor. Soc.*, **98**, 565–588, <https://doi.org/10.1175/BAMS-D-15-00268.1>.
- Blumberg, W. G., K. T. Halbert, T. A. Supinie, P. T. Marsh, R. L. Thompson, and J. A. Hart, 2017: SHARPPy: An open-source sounding analysis toolkit for the atmospheric sciences. *Bull. Amer. Meteor. Soc.*, **98**, 1625–1636, <https://doi.org/10.1175/BAMS-D-15-00309.1>.
- Britt, K. C., P. S. Skinner, P. L. Heinselman, and K. H. Knopfmeier, 2018: Strengths and limitations of NSSL's experimental Warn-on-Forecast system for ensembles in forecasting cyclic mesocyclogenesis. *29th Conf. on Severe Local Storms*, Stowe, VT, Amer. Meteor. Soc., 141, <https://ams.confex.com/ams/29SLS/meetingapp.cgi/Paper/348411>.
- Bröcker, J., and L. A. Smith, 2008: From ensemble forecasts to predictive distribution functions. *Tellus*, **60A**, 663–678, <https://doi.org/10.1111/j.1600-0870.2008.00333.x>.
- Brooks, H. E., J. W. Lee, and J. P. Craven, 2003: The spatial distribution of severe thunderstorm and tornado environments from global reanalysis data. *Atmos. Res.*, **67–68**, 73–94, [https://doi.org/10.1016/S0169-8095\(03\)00045-0](https://doi.org/10.1016/S0169-8095(03)00045-0).
- Buizza, R., M. Milleer, and T. N. Palmer, 1999: Stochastic representation of model uncertainties in the ECMWF ensemble prediction system. *Quart. J. Roy. Meteor. Soc.*, **125**, 2887–2908, <https://doi.org/10.1002/qj.49712556006>.
- Bunkers, M. J., B. A. Klimowski, J. W. Zeitle, R. L. Thompson, and M. L. Weisman, 2000: Predicting supercell motion using a new hodograph technique. *Wea. Forecasting*, **15**, 61–79, [https://doi.org/10.1175/1520-0434\(2000\)015<0061:PSMUAN>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0061:PSMUAN>2.0.CO;2).
- Burgess, D. W., V. T. Wood, and R. A. Brown, 1982: Mesocyclone evolution statistics. Preprints, *12th Conf. on Severe Local Storms*, San Antonio, TX, Amer. Meteor. Soc., 422–424.
- Burke, P. C., and D. M. Schultz, 2004: A 4-yr climatology of cold-season bow echoes over the continental United States. *Wea. Forecasting*, **19**, 1061–1074, <https://doi.org/10.1175/811.1>.
- Cintineo, R. M., and D. J. Stensrud, 2013: On the predictability of supercell thunderstorm evolution. *J. Atmos. Sci.*, **70**, 1993–2011, <https://doi.org/10.1175/JAS-D-12-0166.1>.
- Coffer, B. E., and M. D. Parker, 2018: Is there a “tipping point” between simulated nontornadic and tornadic supercells in VORTEX2 environments? *Mon. Wea. Rev.*, **146**, 2667–2693, <https://doi.org/10.1175/MWR-D-18-0050.1>.
- , —, J. M. L. Dahl, L. J. Wicker, and A. J. Clark, 2017: Volatility of tornadogenesis: An ensemble of simulated nontornadic and tornadic supercells in VORTEX2 environments. *Mon. Wea. Rev.*, **145**, 4605–4625, <https://doi.org/10.1175/MWR-D-17-0152.1>.
- Crook, N. A., 1996: Sensitivity of moist convection forced by boundary layer processes to low-level thermodynamic fields. *Mon. Wea. Rev.*, **124**, 1767–1785, [https://doi.org/10.1175/1520-0493\(1996\)124<1767:SOMCFB>2.0.CO;2](https://doi.org/10.1175/1520-0493(1996)124<1767:SOMCFB>2.0.CO;2).
- Curry, M. J., and R. C. Murty, 1974: Thunderstorm-generated gravity waves. *J. Atmos. Sci.*, **31**, 1402–1408, [https://doi.org/10.1175/1520-0469\(1974\)031<1402:TGGW>2.0.CO;2](https://doi.org/10.1175/1520-0469(1974)031<1402:TGGW>2.0.CO;2).
- Doswell, C. A., H. E. Brooks, and R. A. Maddox, 1996: Flash flood forecasting: An ingredients-based methodology. *Wea. Forecasting*, **11**, 560–581, [https://doi.org/10.1175/1520-0434\(1996\)011<0560:FFFAIB>2.0.CO;2](https://doi.org/10.1175/1520-0434(1996)011<0560:FFFAIB>2.0.CO;2).
- Droegemeier, K. K., and J. Levit, 1993: The sensitivity of numerically-simulated storm evolution to initial conditions. Preprints, *17th Conf. on Severe Local Storms*, St. Louis, MO, Amer. Meteor. Soc., 431–435.
- Duc, L., K. Saito, and H. Seko, 2013: Spatial-temporal fractions verification for high-resolution ensemble forecasts. *Tellus*, **65A**, 18171, <https://doi.org/10.3402/tellusa.v65i0.18171>.

- Durran, D. R., and J. A. Weyn, 2016: Thunderstorms do not get butterflies. *Bull. Amer. Meteor. Soc.*, **97**, 237–243, <https://doi.org/10.1175/BAMS-D-15-00070.1>.
- Elmore, K. L., D. J. Stensrud, and K. C. Crawford, 2002: Explicit cloud-scale models for operational forecasts: A note of caution. *Wea. Forecasting*, **17**, 873–884, [https://doi.org/10.1175/1520-0434\(2002\)017<0873:ECSMFO>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<0873:ECSMFO>2.0.CO;2).
- Flora, M. L., C. K. Potvin, and L. J. Wicker, 2018: Practical predictability of supercells: Exploring ensemble forecast sensitivity to initial condition spread. *Mon. Wea. Rev.*, **146**, 2361–2379, <https://doi.org/10.1175/MWR-D-17-0374.1>.
- Fujita, T. T., 1990: Downbursts: Meteorological features and wind field characteristics. *J. Wind Eng. Ind. Aerodyn.*, **36**, 75–86, [https://doi.org/10.1016/0167-6105\(90\)90294-M](https://doi.org/10.1016/0167-6105(90)90294-M).
- Gilleland, E., D. Ahijevych, B. G. Brown, B. Casati, and E. E. Ebert, 2009: Intercomparison of spatial forecast verification methods. *Wea. Forecasting*, **24**, 1416–1430, <https://doi.org/10.1175/2009WAF2222269.1>.
- , —, —, and E. E. Ebert, 2010: Verifying forecasts spatially. *Bull. Amer. Meteor. Soc.*, **91**, 1365–1373, <https://doi.org/10.1175/2010BAMS2819.1>.
- Gleick, J., 1987: *Chaos: Making A New Science*. Vintage, 352 pp.
- Grams, J. S., R. L. Thompson, D. V. Snively, J. A. Prentice, G. M. Hodges, and L. J. Reames, 2012: A climatology and comparison of parameters for significant tornado events in the United States. *Wea. Forecasting*, **27**, 106–123, <https://doi.org/10.1175/WAF-D-11-00008.1>.
- Grebogi, C., E. Ott, and J. A. Yorke, 1987: Chaos, strange attractors, and fractal basin boundaries in nonlinear dynamics. *Science*, **238**, 632–638, <https://doi.org/10.1126/science.238.4827.632>.
- Herman, G. R., E. R. Nielsen, and R. S. Schumacher, 2018: Probabilistic verification of Storm Prediction Center convective outlooks. *Wea. Forecasting*, **33**, 161–184, <https://doi.org/10.1175/WAF-D-17-0104.1>.
- Hohenegger, C., D. Lüthi, and C. Schär, 2006: Predictability mysteries in cloud-resolving models. *Mon. Wea. Rev.*, **134**, 2095–2107, <https://doi.org/10.1175/MWR3176.1>.
- Igel, A. L., M. R. Igel, and S. C. van den Heever, 2015: Make it a double? Sobering results from simulations using single-moment microphysics schemes. *J. Atmos. Sci.*, **72**, 910–925, <https://doi.org/10.1175/JAS-D-14-0107.1>.
- Jankov, I., W. A. Gallus, M. Segal, B. Shaw, and S. E. Koch, 2005: The impact of different WRF Model physical parameterizations and their interactions on warm season MCS rainfall. *Wea. Forecasting*, **20**, 1048–1060, <https://doi.org/10.1175/WAF888.1>.
- Kain, J. S., and Coauthors, 2008: Some practical considerations regarding horizontal resolution in the first generation of operational convection-allowing NWP. *Wea. Forecasting*, **23**, 931–952, <https://doi.org/10.1175/WAF2007106.1>.
- Kirkpatrick, J. C., E. W. McCaul Jr., and C. Cohen, 2007: The motion of simulated convective storms as a function of basic environmental parameters. *Mon. Wea. Rev.*, **135**, 3033–3051, <https://doi.org/10.1175/MWR3447.1>.
- , —, and —, 2011: Sensitivities of simulated convective storms to environmental CAPE. *Mon. Wea. Rev.*, **139**, 3514–3532, <https://doi.org/10.1175/2011MWR3631.1>.
- Knievel, J. C., G. H. Bryan, and J. P. Hacker, 2007: Explicit numerical diffusion in the WRF Model. *Mon. Wea. Rev.*, **135**, 3808–3824, <https://doi.org/10.1175/2007MWR2100.1>.
- Lawson, J., and W. A. Gallus Jr., 2016a: On contrasting ensemble simulations of two Great Plains bow echoes. *Wea. Forecasting*, **31**, 787–810, <https://doi.org/10.1175/WAF-D-15-0060.1>.
- , and —, 2016b: Adapting the SAL method to evaluate reflectivity forecasts of summer precipitation in the central United States. *Atmos. Sci. Lett.*, **17**, 524–530, <https://doi.org/10.1002/asl.687>.
- , J. S. Kain, N. Yussouf, D. C. Dowell, D. M. Wheatley, K. H. Knopfmeier, and T. A. Jones, 2018: Advancing from convection-allowing NWP to Warn-on-Forecast: Evidence of progress. *Wea. Forecasting*, **33**, 599–607, <https://doi.org/10.1175/WAF-D-17-0145.1>.
- Leroy, S. S., and M. J. Rodwell, 2014: Leveraging highly accurate data in diagnosing errors in atmospheric models. *Bull. Amer. Meteor. Soc.*, **95**, 1227–1233, <https://doi.org/10.1175/BAMS-D-12-00143.1>.
- Lillo, S. P., and D. B. Parsons, 2017: Investigating the dynamics of error growth in ECMWF medium-range forecast busts. *Quart. J. Roy. Meteor. Soc.*, **143**, 1211–1226, <https://doi.org/10.1002/qj.2938>.
- Lilly, D. K., 1986: The structure, energetics and propagation of rotating convective storms. Part II: Helicity and storm stabilization. *J. Atmos. Sci.*, **43**, 126–140, [https://doi.org/10.1175/1520-0469\(1986\)043<0126:TSEAPO>2.0.CO;2](https://doi.org/10.1175/1520-0469(1986)043<0126:TSEAPO>2.0.CO;2).
- , 1990: Numerical prediction of thunderstorms—Has its time come? *Quart. J. Roy. Meteor. Soc.*, **116**, 779–798, <https://doi.org/10.1002/qj.49711649402>.
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, [https://doi.org/10.1175/1520-0469\(1963\)020<0130:DNF>2.0.CO;2](https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2).
- , 1969: The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307, <https://doi.org/10.3402/tellusa.v21i3.10086>.
- , 1995: *The Essence of Chaos*. University of Washington Press, 227 pp.
- Mansell, E. R., C. L. Ziegler, and E. C. Bruning, 2010: Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *J. Atmos. Sci.*, **67**, 171–194, <https://doi.org/10.1175/2009JAS2965.1>.
- Markowski, P., and Y. Richardson, 2010: *Mesoscale Meteorology in Mid-latitudes*. Wiley-Blackwell, 407 pp.
- McCaul, E. W., Jr., and M. L. Weisman, 2001: The sensitivity of simulated supercell structure and intensity to variations in the shapes of environmental buoyancy and shear profiles. *Mon. Wea. Rev.*, **129**, 664–687, [https://doi.org/10.1175/1520-0493\(2001\)129<0664:TSSOSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2001)129<0664:TSSOSS>2.0.CO;2).
- , and C. Cohen, 2002: The impact on simulated storm structure and intensity of variations in the mixed layer and moist layer depths. *Mon. Wea. Rev.*, **130**, 1722–1748, [https://doi.org/10.1175/1520-0493\(2002\)130<1722:TIOSSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1722:TIOSSS>2.0.CO;2).
- McMurdie, L. A., and B. Ancell, 2014: Predictability characteristics of landfalling cyclones along the North American West Coast. *Mon. Wea. Rev.*, **142**, 301–319, <https://doi.org/10.1175/MWR-D-13-00141.1>.
- Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137**, 991–1007, <https://doi.org/10.1175/2008MWR2556.1>.
- Mulder, K. J., and D. M. Schultz, 2015: Climatology, storm morphologies, and environments of tornadoes in the British Isles: 1980–2012. *Mon. Wea. Rev.*, **143**, 2224–2240, <https://doi.org/10.1175/MWR-D-14-00299.1>.
- Novak, D. R., D. R. Bright, and M. J. Brennan, 2008: Operational forecaster uncertainty needs and future roles. *Wea. Forecasting*, **23**, 1069–1084, <https://doi.org/10.1175/2008WAF2222142.1>.

- Nusse, H. E., and J. A. Yorke, 1996: Basins of attraction. *Science*, **271**, 1376–1380, <https://doi.org/10.1126/science.271.5254.1376>.
- Palmer, T. N., 1988: Medium and extended range predictability and stability of the Pacific/North American mode. *Quart. J. Roy. Meteor. Soc.*, **114**, 691–713, <https://doi.org/10.1002/qj.49711448108>.
- , A. Döring, and G. Seregin, 2014: The real butterfly effect. *Nonlinearity*, **27**, R123, <https://doi.org/10.1088/0951-7715/27/9/R123>.
- Potvin, C. K., and M. L. Flora, 2015: Sensitivity of idealized supercell simulations to horizontal grid spacing: Implications for Warn-on-Forecast. *Mon. Wea. Rev.*, **143**, 2998–3024, <https://doi.org/10.1175/MWR-D-14-00416.1>.
- Powers, J. G., and Coauthors, 2017: The Weather Research and Forecasting Model: Overview, system efforts, and future directions. *Bull. Amer. Meteor. Soc.*, **98**, 1717–1737, <https://doi.org/10.1175/BAMS-D-15-00308.1>.
- Reinecke, P., and D. Durran, 2009: Initial-condition sensitivities and the predictability of downslope winds. *J. Atmos. Sci.*, **66**, 3401–3418, <https://doi.org/10.1175/2009JAS3023.1>.
- Roberts, N. M., and H. W. Lean, 2008: Scale-selective verification of rainfall accumulations from high-resolution forecasts of convective events. *Mon. Wea. Rev.*, **136**, 78–97, <https://doi.org/10.1175/2007MWR2123.1>.
- Rodwell, M. J., and Coauthors, 2013: Characteristics of occasional poor medium-range weather forecasts for Europe. *Bull. Amer. Meteor. Soc.*, **94**, 1393–1405, <https://doi.org/10.1175/BAMS-D-12-00099.1>.
- Schertzer, D., and S. Lovejoy, 1985: The dimension and intermittency of atmospheric dynamics. *Turbulent Shear Flows 4*, Springer, 7–33.
- Schultz, D., and P. Schumacher, 1999: The use and misuse of conditional symmetric instability. *Mon. Wea. Rev.*, **127**, 2709–2732, [https://doi.org/10.1175/1520-0493\(1999\)127<2709:TUAMOC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1999)127<2709:TUAMOC>2.0.CO;2).
- Schwartz, C. S., and R. A. Sobash, 2017: Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches: A review and recommendations. *Mon. Wea. Rev.*, **145**, 3397–3418, <https://doi.org/10.1175/MWR-D-16-0400.1>.
- , and Coauthors, 2010: Toward improved convection-allowing ensembles: Model physics sensitivities and optimizing probabilistic guidance with small ensemble membership. *Wea. Forecasting*, **25**, 263–280, <https://doi.org/10.1175/2009WAF2222267.1>.
- Shen, B.-W., R. A. Pielke Sr., X. Zeng, S. Faghieh-Naini, C. L. Shie, J. J. Baik, and T. A. L. Reyes, 2018: Butterfly effects of the first and second kinds: New insights revealed by high-dimensional Lorenz models. *11th Int. Conf. on Chaotic Modeling, Simulation and Applications*, Rome, Italy.
- Sherburn, K. D., and M. D. Parker, 2014: Climatology and ingredients of significant severe convection in high-shear, low-CAPE environments. *Wea. Forecasting*, **29**, 854–877, <https://doi.org/10.1175/WAF-D-13-00041.1>.
- , —, J. R. King, and G. M. Lackmann, 2016: Composite environments of severe and nonsevere high-shear, low-CAPE convective events. *Wea. Forecasting*, **31**, 1899–1927, <https://doi.org/10.1175/WAF-D-16-0086.1>.
- Skinner, P. S., and Coauthors, 2018: Object-based verification of a prototype Warn-on-Forecast system. *Wea. Forecasting*, **33**, 1225–1250, <https://doi.org/10.1175/WAF-D-18-0020.1>.
- Stensrud, D. J., J. W. Bao, and T. T. Warner, 2000: Using initial condition and model physics perturbations in short-range ensemble simulations of mesoscale convective systems. *Mon. Wea. Rev.*, **128**, 2077–2107, [https://doi.org/10.1175/1520-0493\(2000\)128<2077:UICAMP>2.0.CO;2](https://doi.org/10.1175/1520-0493(2000)128<2077:UICAMP>2.0.CO;2).
- Sterk, A. E., D. B. Stephenson, M. P. Holland, and K. R. Mylne, 2016: On the predictability of extremes: Does the butterfly effect ever decrease? *Quart. J. Roy. Meteor. Soc.*, **142**, 58–64, <https://doi.org/10.1002/qj.2627>.
- Tennekes, H., 1978: Turbulent flow in two and three dimensions. *Bull. Amer. Meteor. Soc.*, **59**, 22–28, [https://doi.org/10.1175/1520-0477\(1978\)059<0022:TFITAT>2.0.CO;2](https://doi.org/10.1175/1520-0477(1978)059<0022:TFITAT>2.0.CO;2).
- Trapp, R. J., G. J. Stumpf, and K. L. Manross, 2005: A reassessment of the percentage of tornadic mesocyclones. *Wea. Forecasting*, **20**, 680–687, <https://doi.org/10.1175/WAF864.1>.
- Wandishin, M. S., D. J. Stensrud, S. L. Mullen, and L. J. Wicker, 2008: On the predictability of mesoscale convective systems: Two-dimensional simulations. *Wea. Forecasting*, **23**, 773–785, <https://doi.org/10.1175/2008WAF2007057.1>.
- , —, —, and —, 2010: On the predictability of mesoscale convective systems: Three-dimensional simulations. *Mon. Wea. Rev.*, **138**, 863–885, <https://doi.org/10.1175/2009MWR2961.1>.
- Weisman, M. L., and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110**, 504–520, [https://doi.org/10.1175/1520-0493\(1982\)110<0504:TDONSC>2.0.CO;2](https://doi.org/10.1175/1520-0493(1982)110<0504:TDONSC>2.0.CO;2).
- , and —, 1984: The structure and classification of numerically simulated convective storms in directionally varying wind shears. *Mon. Wea. Rev.*, **112**, 2479–2498, [https://doi.org/10.1175/1520-0493\(1984\)112<2479:TSACON>2.0.CO;2](https://doi.org/10.1175/1520-0493(1984)112<2479:TSACON>2.0.CO;2).
- , and R. Rotunno, 2000: The use of vertical wind shear versus helicity in interpreting supercell dynamics. *J. Atmos. Sci.*, **57**, 1452–1472, [https://doi.org/10.1175/1520-0469\(2000\)057<1452:TUOVWS>2.0.CO;2](https://doi.org/10.1175/1520-0469(2000)057<1452:TUOVWS>2.0.CO;2).
- Wernli, H., M. Paulat, M. Hagen, and C. Frei, 2008: SAL—A novel quality measure for the verification of quantitative precipitation forecasts. *Mon. Wea. Rev.*, **136**, 4470–4487, <https://doi.org/10.1175/2008MWR2415.1>.
- Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast System. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, <https://doi.org/10.1175/WAF-D-15-0043.1>.
- Williams, G. P., 1997: *Chaos Theory Tamed*. Joseph Henry Press, 499 pp.
- Yussouf, N., J. S. Kain, and A. J. Clark, 2016: Short-term probabilistic forecasts of the 31 May 2013 Oklahoma tornado and flash flood event using a continuous-update-cycle storm-scale ensemble system. *Wea. Forecasting*, **31**, 957–983, <https://doi.org/10.1175/WAF-D-15-0160.1>.
- Zhang, F., C. Snyder, and R. Rotunno, 2002: Mesoscale predictability of the “surprise” snowstorm of 24–25 January 2000. *Mon. Wea. Rev.*, **130**, 1617–1632, [https://doi.org/10.1175/1520-0493\(2002\)130<1617:MPOTSS>2.0.CO;2](https://doi.org/10.1175/1520-0493(2002)130<1617:MPOTSS>2.0.CO;2).
- Zhang, Y., F. Zhang, D. J. Stensrud, and Z. Meng, 2016: Intrinsic predictability of the 20 May 2013 tornadic thunderstorm event in Oklahoma at storm scales. *Mon. Wea. Rev.*, **144**, 1273–1298, <https://doi.org/10.1175/MWR-D-15-0105.1>.